

## SEQUENCE LISTING

<110> Cole, Stewart  
 Pym, Alexander S  
 Brosch, Roland  
 Brodin, Priscille  
 Majlessi, Laleh  
 Demangel, Caroline  
 Leclerc, Claude

<120> Identification of virulence associated regions RD1 and  
 RD5 leading to improve vaccine of *M. bovis* BCG and *M.*  
*microti*

<130> D20217

<150> PCT/IB03/01789

<151> 2003-04-01

<150> EP 02/290864

<151> 2002-04-05

<160> 75

<170> PatentIn Ver. 2.1

<210> 1

<211> 31808

<212> DNA

<213> *Mycobacterium tuberculosis*

<220>

<223> Insert of cosmid RD1-2F9 corresponding to sequence  
 in the genome of *mycobacterium tuberculosis* H37Rv

<400> 1

```

gatccgacca caccagcccg gcaccccgcg gggatactcg ccccgctccg cgtccggaga 60
tcgctgcccc gcgccaccgc ctggccggca cgctgctgcc gctacgccac cagggccgcc 120
gcgcctgcct tcagctccac tgcgtccatt gccggaccgc gcttggccac gccagccgga 180
ggccccgcca ccgagcacct gggccgaccc cgccctggcg ccgatacgca gtcggacgcg 240
accgggcgag cgtgggtggc gacgcatggt gcggctgggc acctttggcc ttgtcggcct 300
gggcccggtc ggcattgcagc gccaggaggc ccaattcgaa gcaacgatac gaaccgtcct 360
gcatggcaac cacaaggctc cgtgctggg caaaggagggt gtgggaaaga cgtcgggttc 420
ggcgtgcgtc ggatcgatcc ttgccgaact gcgccagcag gaccgtatcg tcgggatcga 480
cgccgacacc gccttcggca ggctgagcag ccgaatcgat cctcgagcag ctggttcggt 540
ctgggagctg accaccgaca cgaatctgcg gtccttcacc gatatacccg cgcgcctggg 600
ccgaaattcc cggggactgt acgtcctggc aggccagccg gcatccggtc cgcgcggggt 660
gctcgatccg gccatctacc gcgaagccgc cctaagggtg gatcaccatt tcgcaatctc 720
ggtgatcgac tgcggttcct ccatggaggc ggcggtcacc caggaagtat tgcgcgatgt 780
ggatgctctg atcgtggtgt cctcgccctg ggcggtggtt gcctccgctg ccgccaacac 840
catcgaatgg ctgtcggatt atggcctgac aggtttgttg cgacgcagca tcgtggtgct 900
caacgattcg gacggacacg ccgacaagcg caccaagtca ttgctggccc aggaattcat 960
cgaccacggg cagcctgttg tcgaggtgcc cttcgatccc catttgccgc ccgggggggt 1020
catcgatatg agccacgaaa tggccccgac gacgcggctg aaaatcctgc aggtcgccgc 1080
gacggtgacg gcgtacttcg cgtcgcgacc cgccgacgca cacggcagcc cgccccggtg 1140
acctgctgg ctgaccgggt cggcaacagc aggatcgccc gagcgcaggc ctgcaaaacg 1200
tcaatctcgg cgcccatcgt cgaatcctgg cgggcgcaac gcggcgcgca atgtggacag 1260
cgcgagaaat cttgtcgatg ttctcgcgct gtccacatcc agggcatctc accgccactg 1320
ttccgcagac ccctcgaacc agcgggtccg gcggcggttg cgtcatgccg attgggcaga 1380
cacccggtgg tcgcgcaccg ggtaaccgtt gcgctcggcc agggatcgca gctggcccaa 1440
cgcgaatgcc cgcgcccggc ctgattcggg aattacgacc cctgcccaca gcccttcgcg 1500

```

acccgcggac	tcgacggcgt	cgcgtgcaca	cagccaccgg	cgcgggcaag	cccggcacag	1560
ggtcttggcc	tcgtcgtcgg	gagtcgtcgt	ccaacgatcg	ggatcttgcg	tgcaaacgcc	1620
gagcgggacc	tcatacaggg	cggttactgt	catgtctacg	ttcctccaga	aagcgttgca	1680
ggttgtagcc	tctgccgcga	aagcgtatcg	cattaaccat	agcgatgcaa	cagtttcctc	1740
ctctgcctgc	ctagcgggtgc	tgcggctccg	gttcggcgag	ctccgagctc	tagtgcgcg	1800
accgccgagt	accagggcat	agatcctgtt	aatcagctgt	gtatctggcc	tcgccggcgc	1860
gtatccgacc	ccttcgggca	gatcttccag	gaaaagtgtt	ctgacatgcg	acagttcagg	1920
tgtgaagtga	actgtagcgg	cagttcgggt	tggctaggaa	actatttcca	tagcgggccc	1980
tcgcgtcgtc	agatccaaaa	tgtagcgaag	tcatagcagt	agaagggtgc	aacggttagg	2040
atggcgggcg	agcggaaagt	ctgcccaccg	tcccggctag	taccgcgcaa	taagggatca	2100
acgcagatgt	ctaaagcagg	gtcgactgtc	ggaccggcgc	cgctggctgc	gtgcagcggc	2160
ggcacatcag	acgtgattga	gccccgtcgc	ggtgtcgcga	tcattggcca	ctcgtgccga	2220
gtcggcacc	agatcgacga	ttctcgaatc	tctcagacac	atctgcgagc	ggtatccgat	2280
gatggacggt	ggcggatcgt	cggcaacatc	ccgagaggta	tgttcgtcgg	cggacgacgc	2340
ggcagctcgg	tgaccgtcag	cgataagacc	ctaaccgat	tcggcgatcc	ccctggaggc	2400
aaggcgttga	cgttcgaagt	cgtcaggccg	tcggattccg	ctgcacagca	cggccgcgta	2460
caaccatcag	cggacctgtc	ggacgaccgc	gcgcacaacg	ctgcgccggt	cgcaccggac	2520
cccggcgtgg	ttcgcgcagc	ggcggccgcg	gctgcgcgcc	gtcgtgaact	tgacatcagc	2580
caacgcagct	tggcggccga	cgggatcatc	aacgcgggcg	cgctcatcgc	gttcgagaaa	2640
ggccgtagtt	ggccccggga	acggaccgcg	gcaaaactcg	aagaagtgtc	gcagtggccc	2700
gctggaacca	tcgcgcgaat	ccgtcggggc	gagcccaccg	agcccgaac	aaaccccagc	2760
gcgtcccccg	gactccggcc	tgcgcagcgc	ccggcgtcct	tgatcgcgca	ggctgtcacc	2820
gccgccgtag	acggctgcag	tctggctatc	gcagcgttgc	cggcgaccga	ggaccccag	2880
ttcaccgaac	gtgccgcgcc	gatccttgct	gatttgccgc	agctcgaggc	gattgccgtc	2940
caagcaaccc	gcacagccg	gattaccccg	gaattgatca	aggcgttggg	cgcggtacgt	3000
cgccaccacg	acgaattaat	gaggctggga	gcaaccgccc	ctggtgccac	actggcgcag	3060
cgcttatatg	ccgcacggcg	gcgcgcgaac	ctttccaccc	tggagactgc	ccaagcggcc	3120
ggcgctcgag	aagaaatgat	cgtcggcgcc	gaatgcaggg	aagagttgcc	agccgaggcc	3180
accgaagcga	tcgaagcact	gatccgtcag	atcaattgag	gtcggctccg	agcgtcccac	3240
aagtacaggc	acgccgtaac	gctcaagttc	aacggtccgg	ggaacgcgcg	cgttctccgg	3300
cgtttgacgg	tgcgttccat	cgtgccgcga	acttgaaaac	gccagcgtca	ccaaaaaatt	3360
cgtgcaccaa	ccccctccg	agcgtgtgta	agctcaatgt	gcagtgcaaa	ggtgcagata	3420
atgatggcgc	accggaacgg	cgagcgtaa	gaaacacata	aatggcatcg	ggtagcggtc	3480
tttgcaagac	gacgagtaac	tttatttggg	gccagttact	cttgcttggg	gagggaatcc	3540
ccgaccacgg	cgacattttc	aacaccgggt	cgtcgtgtgt	caaacaaatc	agcgacaaaa	3600
tgggactcgc	cattccggcg	accaactgga	tcggccaagc	ggcggaagct	tacctaaacc	3660
agaacatcgc	tgcaacaact	cgcgcacagg	tgatggcgga	tctcgacaaa	ttaaccggca	3720
acatgatctc	gaatcaggcc	aaatacgtct	ccgatacgcg	cgacgtcctg	cgggccatga	3780
agaagatgat	tgacggtgtc	tacaaggttt	gtaagggcct	cgaaaagatt	ccgctgctcg	3840
gccacttggt	gtcgtgggag	ctcgcaatcc	ctatgtccgg	catcgcgatg	gccgttgctg	3900
gcggcgcatt	gctctatcta	acgattatga	cgctgatgaa	tgcgaccaac	ctgaggggaa	3960
ttctcggcag	gctgatcgag	atgttgacga	ccttgccaaa	gttccccggc	ctgcccgggt	4020
tgcccagcct	gcccgcacatc	atcgacggcc	tctggccgcc	gaagtgtccc	gacattccga	4080
tccccggcct	gcccgcacatc	ccgggcctac	ccgacttcaa	atggccgccc	acccccggca	4140
gcccgttggt	ccccgcacctc	ccgtcgttcc	cagggttccc	cgggttcccg	gagttccccg	4200
ccatcccccg	gttccccgca	ctgcccgggt	tgcccagcat	tcccaacttg	ttccccggct	4260
tgccgggtct	gggcgcacctg	ctgcccggcg	taggcgattt	gggcaagtta	cccacctgga	4320
ctgagctggc	cgttttgcc	gacttcttgg	gcggcttcgc	cggcctgccc	agcttggggt	4380
ttggcaatct	gctcagcttt	gccagtttgc	ccaccgtggg	tcaggtgacc	gccaccatgg	4440
gtcagctgca	acagctcgtg	gcggccggcg	gtggccccag	ccaactggcc	agcatgggca	4500
gccaacaagc	gcaactgata	tcgtcgcagg	cccagcaagg	aggccagcag	cacgccaccc	4560
tcgtgagcga	caagaaggaa	gacgaggaag	gcgtggccga	ggcggagcgt	gcacccatcg	4620
acgctggcac	cgcggccagc	caacgggggc	aggaggggac	cgtcctttga	tcggacaccg	4680
agtcgccagc	aggtctgtgc	catagcgagt	cgaagccata	gcgagtagaa	agttaaacgt	4740
agaggagggt	tcaacctatg	accggatttc	tcggtgctgt	gccttcgttc	ctgaagggtgc	4800
tggcgggcat	gcacaacgag	atcgtgggtg	atatcaaaa	ggcgaccgat	acggctcgccg	4860
ggattagcgg	acgagttcag	cttaccatg	gttcgttcac	gtcgaaattc	aatgacacgc	4920
tgcaagagtt	tgagaccacc	cgtagcagca	cgggcacggg	tttgaggga	gtcaccagcg	4980
gactggccaa	taatctgctc	gcagccgcgc	gcgcctacct	caaggccgac	gatggcctag	5040
ccggtgttat	cgacaagatt	ttcggttgat	catgacgggt	ccgtccgctg	caggccgcgc	5100
gggcaccgcc	gacaacgtgg	tcggcgtcga	ggtaaccatc	gacggcatgt	tggtgatcgc	5160

cgatcgggta	cacctgggtg	atttccctgt	cacgcttggt	attcggccga	atatcccgca	5220
agaggatctg	cgagacatcg	tctgggaaca	ggtgcagcgt	gacctcacag	cgcaaggggt	5280
gctcgacctc	cacggggagc	cccaaccgac	ggtcgcgag	atggtcgaaa	ccctgggcag	5340
gccagatcgg	accttggagg	gtcgttggtg	gcggcgcgac	attggcgggc	tcatgggtgc	5400
cttcgtcgtg	tgccgcagg	gcgaccgcca	tgtgatcgcg	gcgcgcgacg	gcgacatgct	5460
ggtgctgcag	ttggtggcgc	cgcagggtcg	cttggcgggc	atggtgacag	cgggtgctgg	5520
gcccgcgcaa	cccgcacaac	tccaacccct	gacgggtgtg	gcaaccgagc	tagccgaatg	5580
cacaaccgcg	tcccaattga	cgcaatacgg	tatcgaccgc	gcctcggccc	gcgtctatgc	5640
cgagatcgtg	ggtaaccgca	ccgggtgggt	ggagatcggt	gccagccaac	gccaccccg	5700
cggcaccacg	acgcagaccg	acgccgcgcg	tggcgctcgt	gactccāagc	tcggtaggct	5760
ggtgtcgctt	ccccgcgctg	ttggaggcga	cctgtacgga	agcttcctgc	ccggcactca	5820
gcagaacttg	gagcgtgcgc	tggacggctt	gctagagctg	ctccctgcgg	gcgcttggt	5880
agatcacacc	tcagatcacg	cacaagcctc	ctcccagggc	tgacccctca	catctccgct	5940
acgacttcag	aaagggagcg	catggtggac	ccgccgggca	acgacgacga	ccacgggtgat	6000
ctcgacgccc	tcgattttct	cgcgcgccac	accaacgagg	cgtcgcgcgt	ggacgcctta	6060
gacgactatg	cgcgggtgca	gaccgatgac	gccgaaggcg	acctggacgc	cctccatgcg	6120
ctcaccgaac	gcgacgagga	gccggagctg	gagttgttca	cgggtgaccaa	ccctcaagg	6180
tcggtgtcgg	tctcaaccct	gatggacggc	agaatccagc	acgtcgagct	gacggacaa	6240
gcgaccagca	tgtccgaagc	gcagctggcc	gacgagatct	tcgttattgc	cgatctggcc	6300
cgccaaaagg	cgcgggcgtc	gcagtacacg	ttcatggtgg	agaacatcgg	tgaactgacc	6360
gacgaagacg	cagaaggcag	cgccttgctg	cggaatttcg	tggggatgac	cctgaatctg	6420
ccgacgcggg	aagaggctgc	cgcagccgaa	gccgaagtgt	tcgccacccg	ctacgatgtc	6480
gactacacct	cccgggtaca	ggccgatgac	tgatcgcttg	gccagtctgt	tcgaaagcgc	6540
cgtcagcatg	ttgccgatgt	cggaggcgcg	gtcgttagat	ctgttcaccg	agatcaccaa	6600
ctacgacgaa	tccgcttgcg	acgcatggat	cggccggatc	cgggtgtggg	acaccgaccg	6660
ggtgacgctg	tttcgcgcct	ggtattcgcg	ccgcaatttc	ggacagttgt	cgggatcggg	6720
ccagatctcg	atgagcacgt	taaacgccag	gattgccatc	ggggggctgt	acggcgatat	6780
cacctaaccg	gtcacctcgc	cgtagcgat	caccatgggc	tttgccgcgt	gcgaggcagc	6840
gcaaggcaat	tacggcgacg	ccatggaggc	cttagaggcc	gccccggctg	cgggttccga	6900
gcacctgggt	gcgtggatga	aggcggttgt	ctacggcgcg	gccgaacgct	ggaccgacgt	6960
gatcgaccag	gtcaagagtg	ctgggaaatg	gccggacaag	tttttgcccg	gcgcggccgg	7020
tgtggcgcac	ggggttgccg	cggcaaactc	ggccttgctc	accgaagccg	aacgccgact	7080
caccgaggcc	aacgactcgc	cgcgcgggtg	ggcgtgtgcg	cgcgccatcg	cctgggtatct	7140
ggcgatggca	cggcgcgacc	agggcaacga	aagcgccgcg	gtggcgctgc	tggaatgggt	7200
acagaccact	cacccccgagc	ccaaagtggc	tgccggcgct	aaggatccct	cctaccgggt	7260
gaagacgacc	accgccgaac	agatcgcatc	ccgcgccgat	ccctgggatc	cgggcagtgt	7320
cgtgaccgac	aactccggcc	gggagcggct	gctcgccgag	gcccgaagcc	aactcgaccg	7380
ccaaattggg	ctcaaccggg	ttaaaaatca	gattgaacgc	taccgcgcgg	cgacgctgat	7440
ggccccgggt	cgcgcgccca	agggtatgaa	ggtcgcccag	cccagcaagc	acatgatctt	7500
caccggaccg	cccgggtaccg	gcaagaccac	gatcgcgcg	gtggtggcca	atatcctggc	7560
cggcttaggc	gtcattgccc	aacccaaact	cgtcgagacg	tcgcgcaagg	acttcgctgc	7620
cgagtacgag	gggcaatcgg	cgggtcaagac	cgctaagacg	atcgatcagg	cgctgggcgg	7680
ggtgcttttc	atcgacgagg	cttatgcgct	ggtgcaggaa	agagacggcc	gcaccgatcc	7740
gttcgggtcaa	gaggcgctgg	acacgctgct	ggcgcggtat	gagaacgacc	gggaccgggt	7800
ggtggtgatc	atcgccgggt	acagctccga	catagatcgg	ctgctggaaa	ccaacgagg	7860
tctgcggtcg	cggttcgcca	ctcgcatcga	gttcgacacc	tattcccccg	aggaactcct	7920
cgagatcgcc	aacgtcattg	ccgctgctga	tgattcgggc	ttgaccgcag	aggcgccgca	7980
gaactttctt	caggccgcca	agcagttgga	gcagcgcatg	ttgcgcggcc	ggcgcgccct	8040
ggacgtcgcc	ggcaacgggt	ggtatgcgcg	ccagctgggt	gaggccagcg	agcaatgccg	8100
ggacatgcgt	ctagcccagg	tcctcgatat	cgacaccctc	gacgaagacc	ggcttcgcga	8160
gatcaacggc	tcagatatgg	cggaggctat	cgcgcgggtg	cacgcacacc	tcaacatgag	8220
agaatgaact	atggggcttc	gcctcaccac	caaggttcag	gttagcgggt	ggcgttttct	8280
gctgcgccgg	ctcgaacacg	ccatcgctgc	ccgggacacc	cggatgtttg	acgaccgcgt	8340
gcagttctac	agccgctcga	tcgctcttgg	catcgctcgt	gcggtcctga	ttctggcggg	8400
tgcgcgcgtg	ctggcgctact	tcaaaccaca	aggcaaacct	ggcggcacca	gcctgttcac	8460
cgaccgcgcg	accaaccagc	tttacgtgct	gctgtccgga	cagttgcata	cgggtctaaa	8520
cctgacttcg	gcgcggctgg	tgctgggcaa	tcgggccaac	ccggccaccg	tgaagtcctc	8580
cgaactgagc	aagctgccga	tgggcccagac	cggttggaat	cccggcgccc	cctacgccac	8640
gcctgttttc	gcgggagca	cctcgatctg	gacctatgc	gacaccgtcg	cccgagccga	8700
ctccacttcc	ccggtagtgc	agaccgcggt	catcgcgatg	ccggtggaga	tcgatgcttc	8760
gatcgatccg	ctccagtcac	acgaagcggt	gctggtgtcc	taccagggcg	aaacctggat	8820

cgtcacaact	aaggggacgcc	acgccataga	tctgaccgac	cgcgccctca	cctcgtcgat	8880
ggggataccg	gtgacggcca	ggccaacccc	gatctcggag	ggcatgttca	acgcgctgcc	8940
tgatatgggg	ccctggcagc	tgccgccgat	accggcggcg	ggcgcgccca	attcgcttgg	9000
cctacctgat	gatctagtga	tccggtcggg	cttcagatc	cacaccgaca	agggcccgcga	9060
atactatgtg	gtgctgcccc	acggcatcgc	gcagggtcaac	gcgacaaccg	ctgcggcgct	9120
gcgcgccacc	caggcgacag	ggctggtcgc	gccaccggca	atgggtgccc	gtctgggtcgt	9180
cagaatcgcc	gaacgggtat	acccctcacc	gctacccgat	gaaccgctca	agatcggtgc	9240
ccggccgcag	gatcccgcg	tgtgctgggc	atggcaacgc	agcgccggcg	accagtcgcc	9300
gcagtcaacg	gtgctgtccg	gccggcatct	gccgatatcg	ccctcagcga	tgaacatggg	9360
gatcaagcag	atccacggga	cgggcaccgt	ttaccctcgac	ggcggaataat	tcgtggcact	9420
gcaatcccc	gatcctcgat	acaccgaatc	gatgtactac	atcgatccac	agggcggtgcg	9480
ttatgggggtg	cctaacgcgg	agacagccaa	gtcgctgggc	ctgagttcac	cccaaacgc	9540
gcccctgggag	atcgcttcgtc	tccctggtcga	cggtccgggtg	ctgtcgaaag	atgccgcact	9600
gctcgagcac	gacacgctgc	ccgctgaccc	tagccccga	aaagttccc	ccggagcctc	9660
cggagcccc	tgatgacgac	caagaagttc	actcccacca	ttaccctgtg	ccccgggtg	9720
accccgggcg	agatcagcct	cacgccgccc	gatgacctgg	gcacgcacat	cccaccgtcg	9780
ggcgtccaaa	agatccttcc	ctacgtgatg	ggtggcgcca	tgctcggcat	gàtcgccatc	9840
atggtggccg	gcggcaccag	gcagctgtcg	ccgtacatgt	tgatgatgcc	gctgatgatg	9900
atcgtgatga	tggtcggcgg	tctggccggg	agcaccgggtg	gtggcggcaa	gaaggtgccc	9960
gaaatcaacg	ccgaccgcaa	ggagtacctg	cggtatttgg	caggactacg	cacccgagtgc	10020
acgtcctcgg	ccacctctca	gggtggcggtc	ttctcctacc	acgcaccgca	tcccagggat	10080
ctgttgctga	tcgtcggcac	ccaacggcag	tggtcccggc	cgccaacgc	cgacttctat	10140
gcggccaccc	gaatcggtat	cggtgaccag	ccggcggtgg	atcgattatt	gaagccggcc	10200
gtcggcgggg	agttggccgc	cgccagcgca	gcacctcagc	cgttcctgga	gccgggtcagt	10260
catatgtggg	tggtcaagtt	tctacgaacc	catggattga	tccatgactg	cccgaactg	10320
ctgcaactcc	gtacctttcc	gactatcgcg	atcggcgggg	acttggcggg	ggcagccggc	10380
ctgatgacgg	cgatgatctg	tcacctagcc	gtgttcacc	caccggacct	gctgcagatc	10440
cgggtgctca	ccgaggaacc	cgacgacccc	gactgttcc	ggctcaaagt	gcttccgcac	10500
gtacagcacc	agaccgaaac	cgatgcggcc	gggtccaccc	ggctgatctt	cacgcgccag	10560
gaaggtctgt	cggacctggc	cgcgcgcggg	ccacacgcac	ccgattcgct	tcccggcggc	10620
ccctacgtag	tcgtcgtcga	cctgaccggc	ggcaaggctg	gattcccgc	cgacggtagg	10680
gccggtgtca	cggtgatcac	gttgggcaac	catcgcggt	cggcctaccg	catcagggtg	10740
cacgaggatg	ggacggctga	tgaccggctc	cctaaccaat	cgtttcgcca	ggtgacatcg	10800
gtcaccgatc	ggatgtcgcc	gcagcaagcc	agccgtatcg	cgcgaaaagt	ggccggatgg	10860
tccatcacgg	gcaccatcct	cgacaagacg	tcgcgggtcc	agaagaagg	ggccaccgac	10920
tggcaccagc	tggtcgggtg	gcaaagtgtc	gaggagataa	caccttccc	ctggaggatg	10980
tacaccgaca	ccgaccgtga	ccggctaaag	atcccgtttg	gtcatgaact	aaagaccggc	11040
aacgtcatgt	acctggacat	caaagagggc	gcggaattcg	gcgcgggacc	gcacggcatg	11100
ctcatcgga	ccacggggtc	tggggaagtcc	gaattcctgc	gcacctgat	cctgtcgctg	11160
gtggcaatga	ctcatccaga	tcaggtgaat	ctcctgtc	ccgacttcaa	aggtgggtca	11220
accttctctg	gaatggaaaa	gcttccgcac	actgccgctg	tcgtcaccaa	catggccgag	11280
gaagccgagc	tcgtcagccg	gatgggcgag	gtgttgaccg	gagaactcga	tcggcgccag	11340
tcgatcctcc	gacaggccgg	gatgaaagtc	ggcgcgggcg	gagccctgtc	cggcgtggcc	11400
gaatacagaga	agtaccgcga	acgcgggtgc	gacctacccc	cgctgccaac	gcttttcgtc	11460
gtcgtcgacg	agttcgccga	gctgttgag	agtcacccgg	acttcacog	gctgttcgac	11520
cggatctgcc	gcgtcgggcg	gtcgctgagg	gtccatctgc	tgctggctac	ccagtcgctg	11580
cagaccggcg	gtgttcgcat	cgacaaactg	gagccaaacc	tgacatatcg	aatcgcatg	11640
cgcaccacca	gctctcatga	atccaaggcg	gtaatcgga	caccggaggc	gcagtacatc	11700
accaacaagg	agagcggtgt	cgggtttctc	cgggtcggca	tggaagacc	ggtcaagttc	11760
agcaccttct	acatcagtg	gccatacatg	ccgcggcgcg	caggcgctga	aaccaatggt	11820
gaagccggag	ggcccgggtc	acagaccact	agacaagccg	cgcgcattea	caggttcacc	11880
gcggcaccgg	ttctcgagga	ggcgccgaca	ccgtgaccgc	cgccggcgac	gatgcaaagc	11940
gcagcgatga	ggaggagcgg	cgccaacggc	ccgcgcggcg	gacgatgcaa	agcgacgca	12000
tgaggaggag	cggcgcgcat	gactgtgaa	ccggaagtac	ggacgctgcg	cgaggttggtg	12060
ctggaccagc	tcggcactgc	tgaatcgcg	tgtaataaga	tggtggctgc	gccgttgacc	12120
aatccgggtcc	cgctcaacga	gctcatcgcc	cgtgatcggc	gacaacccct	gcgatttgcc	12180
ctgggggatca	tggtgaacc	gcgcgcgccat	ctacaggatg	tggtggggcg	agacgtttcc	12240
ggggccggcg	gcaacatcgg	tattgggggc	gcacctcaaa	ccgggaagtc	gacgctactg	12300
cagacgatgg	tgatgtcggc	cgccgccaca	cactcacccg	gcaacgttca	gttctattgc	12360
atcgacctag	gtggcgggcg	gctgatctat	ctcgaaaacc	ttccacacgt	cggtgggggtg	12420
gccaatcggt	ccgagccccga	caaggtcaac	cgggtggctg	cagagatgca	agccgtcatg	12480

cggaacggg	aaaccacctt	caaggaacac	cgagtgggct	cgatcgggat	gtaccggcag	12540
ctgcgtgacg	atccaagtca	accggttgcg	tccgatccat	acggcgacgt	ctttctgate	12600
atcgacggat	ggccccggtt	tgctggcgag	ttccccgacc	ttgaggggca	ggttcaagat	12660
ctggccgccc	aggggctggc	gttcggcgctc	cacgtcatca	tctccacgcc	acgctggaca	12720
gagctgaagt	cgcggtgttcg	cgactacctc	ggcaccaaga	tcgagttccg	gcttggtgac	12780
gtcaatgaaa	cccagatcga	ccggattacc	cgcgagatcc	cggcgaatcg	tccgggtcgg	12840
gcagtgtcga	tggaaaagca	ccatctgatg	atcggcgctgc	ccaggttcga	cggcggtgcac	12900
agcgccgata	acctggtgga	ggcgatcacc	gcggggggtga	cgcagatcgc	ttcccagcac	12960
accgaacagg	cacctccggt	gcggttcctg	ccggagcgta	tccacctgca	cgaactcgac	13020
ccgaacccgc	cgggaccaga	gtccgactac	cgcactcgct	gggagattcc	gatcggcttg	13080
cgcgagacgg	acctgacgcc	ggctcactgc	cacatgcaca	cgaacccgca	cctactgate	13140
ttcgggtcgg	ccaaatcggg	caagacgacc	attgccacag	cgatcgcgcg	cgccatttgt	13200
gcccgaacaa	gtccccagca	ggtgcgggtt	atgctcgcgg	actaccgctc	gggcctgctg	13260
gacgcgggtg	cggacaccca	tctgctgggc	gccggcgcg	tcaaccgcaa	cagcgcgctg	13320
ctagacgagg	ccgttcaagc	actggcggtc	aacctgaaga	agcggttgcc	gccgaccgac	13380
ctgacgacgg	cgcagctacg	ctcgcgttcg	tggtggagcg	gatttgacgt	cgtgcttctg	13440
gtcgacgatt	ggcacatgat	cgtgggtgcc	gccgggggga	tgccgccgat	ggcaccgctg	13500
gccccgttat	tgccggcggc	ggcagatatc	gggttgacac	tcattgtcac	ctgtcagatg	13560
agccaggctt	acaaggcaac	catggacaag	ttcgtcggcg	ccgcattcgg	gtcgggcgct	13620
ccgacaattgt	tcctttcggg	cgagaagcag	gaattcccat	ccagtgaagt	caaggtcaag	13680
cggcgccccc	ctggccaggc	atttctcgtc	tcgccagacg	gcaaagaggt	catccaggcc	13740
ccctacatcg	agcctccaga	agaagtgttc	gcagcacccc	caagcgccgg	ttaagattat	13800
ttcattgccg	gtgtagcagg	acccgagctc	agcccggtaa	tcgagttcgg	gcaatgctga	13860
ccatcgggtt	tgtttccggc	tataaccgaa	cggtttgtgt	acgggatata	aatacaggga	13920
gggaagaagt	aggcaaatgg	aaaaaatgtc	acatgatccg	atcgctgccc	acattggcac	13980
gcaagtgagc	gacaacgctc	tgcacggcgt	gacggccggc	tcgacggcgc	tgacgtcggt	14040
gaccgggctg	gttccccggg	gggccgatga	ggtctccgcc	caagcggcga	cggcggttcac	14100
atcggagggc	atccaattgc	tggttccaa	tgcatcgccc	caagaccagc	tccaccgtgc	14160
gggcgaagcg	gtccaggagc	tcgcccgcac	ctattcgcac	atcgacgacg	gcgcgcggcg	14220
cgtcttcgcc	gaataggccc	ccaacacatc	ggagggagtg	atcaccatgc	tgtggcacgc	14280
aatgccaccg	gagctaaata	ccgcacggct	gatggccggc	gcgggtccgg	ctccaatgct	14340
tgccggcgcc	gcgggatggc	agacgcttct	ggcggtctct	gacgctcagg	ccgtcgagtt	14400
gaccgcgcgc	ctgaactctc	tgggagaagc	ctggactgga	ggtggcagcg	acaaggcgct	14460
tgccgctgca	acgccgatgg	tggtctggct	acaaaccgcg	tcaacacagg	ccaagaccgc	14520
tgcgatgcag	gcgacggcgc	aagccgcggc	atacacccag	gccatggcca	cgacgccgtc	14580
gctgccggag	atcgccgcca	accacatcac	ccaggccgtc	cttacggcca	ccaacttctt	14640
cggtatcaac	acgatcccga	tcgcgttgac	cgagatggat	tatttcatcc	gtatgtggaa	14700
ccaggcagcc	ctggcaattg	aggtctacca	ggccgagacc	gcggttaaca	cgcttttcga	14760
gaagctcgag	ccgatggcgt	cgatccttga	tcccggcgcg	agccagagca	cgacgaaccc	14820
gatcttcgga	atgccctccc	ctggcagctc	aacaccgggt	ggccagttgc	cgccggcggc	14880
taccagacc	ctcgccaac	tggttgagat	gagcgcccg	atgcagcagc	tgaccagacc	14940
gctgcagcag	gtgacgtcgt	tggtcagcca	ggtgggcggc	accggcgggc	gcaaccagc	15000
cgacgaggaa	gccgcgcaga	tgggcctgct	cggcaccagt	ccgctgtcga	accatccgct	15060
ggctggtgga	tcaggcccca	gcgcgggcgc	gggcctgctg	cgcgcggagt	cgctacctgg	15120
cgcaggtggg	tcgttgaccc	gcacgcgcgt	gatgtctcag	ctgatcgaaa	agccggttgc	15180
ccctcgggtg	atgccggcgg	ctgctgccgg	atcgtcggcg	acgggtggcg	ccgctccggt	15240
gggtgcggga	gcgatgggccc	agggtgcgca	atccggcggc	tccaccaggc	cgggtctggt	15300
cgcgcgggca	ccgctcgcg	aggagcgtga	agaagacgac	gaggacgact	gggacgaaga	15360
ggacgactgg	tgagctcccg	taatgacaac	agacttcccg	gccacccggg	ccggaagact	15420
tgccaacatt	ttggcgagga	aggtaaagag	agaaagtagt	ccagcatggc	agagatgaag	15480
accgatgccg	ctacctcgc	gcaggaggca	ggtaatttcg	agcggatctc	cggcgacctg	15540
aaaaccacaga	tcgaccaggt	ggagtgcagc	gcaggttcgt	tcgagggcca	gtggcgcggc	15600
gcggcgggga	cggccgccc	ggccgcgggtg	gtgcgcttcc	aagaagcagc	caataagcag	15660
aagcaggaac	tcgacgagat	ctcgacgaat	attcgtcagg	ccggcggtcca	atactcgagg	15720
gccgacgagg	agcagcagca	ggcgctgtcc	tcgcaaatgg	gcttctgacc	cgctaatacg	15780
aaaagaacacg	gagcaaaaa	atgacagagc	agcagtgga	tttcgcgggt	atcgaggccg	15840
cggcaagcgc	aatccaggga	aatgtcacgt	ccattcattc	cctccttgac	gaggggaagc	15900
agtccctgac	caagctcgca	gcggcctggg	gcggtagcgg	ttcgaggcgg	taccagggtg	15960
tccagcaaaa	atgggacgcc	acggctaccg	agctgaacaa	cgcgctgcag	aacctggcgc	16020
ggacgatcag	cgaagccggt	caggcaatgg	cttcgaccga	aggcaacgct	actgggatgt	16080
tcgcataggg	caacgccgag	ttcgcgtaga	atagcgaaac	acgggatcgg	gcgagttcga	16140

ccttcgcgtcg	gtctcgcct	ttctcgtggt	tatacgtttg	agcgcactct	gagaggttgt	16200
catggcgcc	gactacgaca	agctcttcg	gccgcacgaa	ggtatggaag	ctccggacga	16260
tatggcagcg	cagccgttct	tcgacccag	tgcttcgttt	ccgccggcgc	ccgcacggc	16320
aaacctaccg	aagcccaacg	gccagactcc	gcccccgacg	tcgcacgacc	tgctcgagcg	16380
gttcgtgtcg	gccccgcgc	cgccaccccc	acccccacct	ccgcctccgc	caactccgat	16440
gccgatcgcc	gcaggagagc	cgccctcgcc	ggaaccggcc	gcatactaaac	caccacacac	16500
ccccatgccc	atcgccggac	ccgaaccggc	cccacccaaa	ccaccacac	ccccatgcc	16560
catcgccgga	cccgaaccgg	ccccacccaa	accacccaCa	cctccgatgc	ccatcgccgg	16620
acctgcaccc	accccaaccg	aatcccagtt	ggcgcccccc	agaccaccga	caccacaaac	16680
gccaaccgga	gcgcgcgacg	aaccggaatc	accggcgccc	cacgtaccct	cgacggggcc	16740
acatcaaccc	cgccgcacccg	caccagcacc	gccctgggca	aagatgccaa	tcggcgaaac	16800
ccgcgccgct	ccgtccagac	cgtctgcgtc	ccggcgccgaa	ccaccgaccc	ggcctgcccc	16860
ccaacactcc	cgacgtgcgc	gccggggtca	ccgctatcgc	acagacaccg	aacgaaacgt	16920
cggaaggtta	gcaactggtc	catccatcca	ggcgcggtcg	cgggcagagg	aagcatccgg	16980
cgcgacgctc	gcccccgga	cgagccctc	gccagcgccg	ttgggccaac	cgagatcgta	17040
tctggctccg	cccacccgcc	ccgcgcgcgac	agaacctccc	cccagccccct	cgccgcagcg	17100
caactccggt	cggcgtgccg	agcgacgcgt	ccaccccgat	ttagccgccc	aacatgccgc	17160
ggcgcaacct	gattcaatta	cgcccgcaac	cactggcggt	cgtcgccgca	agcgtgcagc	17220
gccgatctc	gacgcgacac	agaaatcctt	aaggccggcg	gccaaggggc	cgaaggtgaa	17280
gaaggtgaag	ccccagaaac	cgaagggcac	gaagccgccc	aaagtgggtg	cgacgcgcg	17340
ctggcgacat	tgggtgcatg	cgttgacgcg	aatcaacctg	ggcctgtcac	ccgacgagaa	17400
gtacgagctg	gacctgcacg	ctcgagtccg	ccgcaatccc	cgcggtcgt	atcagatcgc	17460
cgctcgtcgt	ctcaaagggtg	gggctggcaa	aaccacgctg	acagcagcgt	tggggtcgac	17520
gttggtcag	gtgcgggccc	accggatcct	ggctctagac	gcggatccag	gcgccgga	17580
cctcgccgat	cggttagggc	gacaatccgg	cgcgaccatc	gctgatgtgc	ttgcagaaaa	17640
agagctgtcg	cactacaacg	acatccgcgc	acacactagc	gtcaatgcgg	tcaatctgga	17700
agtgtgcgc	gcaccggaat	acagctcggc	gcagcgcgcg	ctcagcgacg	ccgactggca	17760
tttcctgcgc	gatcctgcgt	cgaggtttta	caacctcgtc	ttgggtgatt	gtggggccgg	17820
cttcttcgac	ccgctgaccc	gcggcggtgt	gtccacgggtg	tcgggtgtcg	tggctgtggc	17880
aagtgtctca	atcgacggcg	cacaacaggc	gtcggtcgcg	ttggactgg	tgcgcaacaa	17940
cggttaccaa	gatttgccga	gccgcgcgatg	cgtggtcatc	aatcacatca	tgccgggaga	18000
acccaatgtc	gcagttaaag	acctggtgcg	gcatttcgaa	cagcaagttc	aaccgggccg	18060
ggctcgtggc	atgccgtggg	acaggcacat	tgcggccgga	accgagattt	cactcgactt	18120
gctcgacct	atctacaagc	gcaaggtcct	cgaattggcc	gcagcgctat	ccgacgattt	18180
cgagaggggt	ggacgtcggt	gagcgcacct	gctgttgctg	ctggtcctac	cgccgcgggg	18240
gcaaccgctg	cgcgccctgc	caccacccgg	gtgacgatcc	tgaccggcag	acggatgacc	18300
gatttggtac	tgccagcgcg	ggtgcccgatg	gaaacttata	ttgacgacac	cgtcgcgggtg	18360
ctttccgagg	tgttggaaga	cacgcccgtc	gatgtactcg	gcggcttcga	ctttaccgcg	18420
caaggcggtg	gggcgttcgc	tcgtcccggga	tcgccgcgcg	tgaagctcga	ccagtcactc	18480
gatgacgccg	gggtggtcga	cgggtcactg	ctgactctgg	tgtcagtcag	tcgcaccgag	18540
cgctaccgac	cgttggtcga	ggatgtcatc	gacgcgatcg	cgtgcttga	cgagtcacct	18600
gagttcgacc	gcacggcatt	gaatcgcttt	gtggggcgcg	cgatcccgt	tttgaccgcg	18660
cccgctcatcg	ggatggcgat	gcgggcgtgg	tgggaaactg	ggcgtagctt	gtgggtggccg	18720
ttggcgattg	gcatcctggg	gatecgtgtg	ctggtaggca	gcttcgtcgc	gaacaggttc	18780
taccagagcg	gccacctggc	cgagtgccta	ctggtcacga	cgtatctgct	gatcgcaacc	18840
gcccgcgcgc	tggccgtgcc	gttgccgcgc	ggggtcaact	cgttgggggc	gccacaagtt	18900
gcggcgccg	ctacggccgt	gctgtttttg	accttgatga	cgcgggggcg	ccctcggaag	18960
cgctcatgag	tggcgtcggt	tgccgtgatc	accgctatcg	cggtcatcgc	ggccgcgcgt	19020
gccttcgggt	atggatacca	ggactgggtc	ccgcgggggg	ggatcgcat	cggtgtgttc	19080
attgtgacga	atgcggccaa	gctgaccgtc	gcggtcgcgc	ggatcgcgct	gccgccgatt	19140
ccggtacccg	gcgaaaccgt	ggacaacgag	gagttgctcg	atcccgtcgc	gaccccgag	19200
gctaccagcg	aagaaacccc	gacctggcag	gccatcatcg	cgtcggtgcc	cgcgctccgcg	19260
gtccggctca	ccgagcgcag	caaactggcc	aagcaacttc	tgatcggata	cgtcacgtcg	19320
ggcaccctga	ttctggctgc	cggtgccatc	gcggctgtgg	tgcgcgggca	cttctttgta	19380
cacagcctgg	tggtcgcggg	tttgatcacg	accgtctgcg	gatttcgctc	gcggctttac	19440
ccgagcgctg	ggtgtgcgtg	ggcggtgctg	gcggcgacgg	tcgcgattcc	gacgggtctg	19500
acggccaaac	tcatcatctg	gtaccgcac	tatgcctggc	tgttggtgag	cgtctacctc	19560
acggtagccc	tggttgcgct	cgtgggtggc	gggtcgatgg	ctcacgtccg	gcgcgtttca	19620
ccggctgtaa	aacgaactct	ggaattgatc	gacggcgcca	tgatcgctgc	catcattccc	19680
atgctgctgt	ggatcaccgg	ggtgtacgac	acggtcgcga	atatccggtt	ctgagccgga	19740
tcggctgatt	ggcggttcct	gacagaacat	cgaggacacg	gcgcaggttt	gcataccttc	19800



ggcgccccgac aaattgctgc gattgagcgt gtggcgcgtc cggtaaaatt tgctcgatgg 19860  
 ggaacacgta taggagatcc ggcaatggct gaaccggttg ccgtcgatcc caccggcttg 19920  
 agcgacggcg ccgcgaaatt ggccggcctc gtttttccgc agcctccggc gccgatcgcg 19980  
 gtcagcggaa cggattcggg ggtagcagca atcaacgaga ccatgccaaag catcgaatcg 20040  
 ctggtcagtg acgggctgcc cggcgtgaaa gccgcctga ctcgaacagc atccaacatg 20100  
 aacgcggcgg cggacgtcta tgcgaagacc gatcagtcac tgggaaccag tttgagccag 20160  
 tatgcattcg gctcgtcggg cgaaggcctg gctggcgctg cctcggtcgg tggtcagcca 20220  
 agtcaggcta cccagctgct gagcacaccc gtgtcacagg tcacgaccca gctcggcgag 20280  
 acggcgcgtg agctggcacc ccgtgttggt gcgacgggtg cgcaactcgt tcagctgggt 20340  
 ccgcacggcg ttcagatgac gcaaaacgca tccccatcg ctcagacgat cagtcaaacc 20400  
 gcccaacag cgcgccagag cgcgcaggcg ggcagcggtc caatgcccg acagcttgcc 20460  
 agcgctgaaa aaccggccac cgagcaagcg gagcgggtcc acgaagtgc aaacgacgat 20520  
 cagggcgacc agggcgacgt gcagccggcg gaggtcgttg ccgcggcacg tgacgaaggc 20580  
 gccggcgcat caccgggcca gcagcccgcg gggggcggtt ccgcgcaagc catggatacc 20640  
 ggagccgggt cccgccagc ggcgagtcgg ctggcgggcc ccgtcgatcc gtcgactccg 20700  
 gcacctcaa caaccacaac gttgtagacc gggcctgcc a ggggtccgt ctcgcacgca 20760  
 gcgcctgttg ctgtcctggc ctcgtcagca tgcggcgggc agggcccggt cgagcaaccc 20820  
 ggtgacgtat tgccagtaca gccagtcgc gacggccaca cgctggacgg ccgcgtcagt 20880  
 cgcagtgtgc gcttggtgca gggcaatctc ctgtgagtgg gcagcgtagg cccggaacgc 20940  
 ccgcagatga gggcctcgc ggcggtagc ggtgtcgttc atgggcttca tcagctcgaa 21000  
 ccacagcatg tgccgtcat cgcccggttg attgacatcc accggcgccg gcggcaacaa 21060  
 gtcagcaaaa cgctgatcgg tagtgtcggc cagctgagcc gccgccgagg ggtcgacgac 21120  
 ctccagccgc gaccggcccg tcattttgcc gctctccgga atgtcatctg gctccagcac 21180  
 aatcttgccc acaccgggat ccgaactggc caactgctcc gcggtaccga tcaccgcccg 21240  
 cagcgtcatg tcgtggaag ccgcccaggc ttgcacggcc aaaaccgggt aggtggcaca 21300  
 gcgtgcaatt tcgtcaaccg ggattgcgtg atccgcgctg gccaaagtaca ccttattcgg 21360  
 caattccatc ccgtcgggta tgtaggccag cccatagctg ttggccacga cgatggaacc 21420  
 gtcggtggte accgcggtga tccagaagaa ccgtagtcg cccgcgttgt tgcggagcg 21480  
 gttgagcgcc gccgcgatgc gtcgcgcaa ccgcagcga tcaccgcggc cagctggcg 21540  
 ggcgctggca gctgcagtgg cggcgctcgg tgccgccga gccgccgaca ccgggatcat 21600  
 cgacaccggc gtaccgtcat ctgcagactc gctgcgatcg ggtttgtcga tgtgatcgg 21660  
 cgacggcggg cgggcaggag gtgccgtccg cgccgaggcc gcccgctgc tcggtgccg 21720  
 cgcttgtcc gaggtagcca ccggcgccc cccagtggca gcatgcgacc ccgcgcccga 21780  
 ggccgcggcc gtacccacgc tcgaacgcgc gcccgctccc acggcggtac cgctcggcg 21840  
 ggccggccgc gcccggtgc ccgggacacc ggacgccga gccggcgta ccgacgcgg 21900  
 ggattcgtcc gcattggcag gccccgactg cgtccccccg cccgcgtgt ggccggcac 21960  
 accaggttgc tccgccaacg ccgcgggtt gacgtgcgg gccggctgc cccctgggt 22020  
 gcccggtgt gctggaccag acggaccgg agtggccgg gtaaccggct gggggccagg 22080  
 cgatggcgcc ggtgcggag ccggctcgg gtgtggagcg ggagctggg taaccggcg 22140  
 ggccgggggt gccgggtgtg ccggggcgac cgggggggtg accggcgta tcgggggtg 22200  
 ctgcctggt gtgcccgtt tgaccgggt caccgggtg accggcttg ccggggtcac 22260  
 cggcgtgac ggagtgcgg gcgttggtg gatcggagt accggcgct ccgggatgg 22320  
 tgtgattgg gttcccggg tgatcgggt tcccggggt atcggggtt ccgggtgccc 22380  
 cgggtgtccc ggggatggca cgaccagggt aggcacgtt ggggggtgg gcgacttctg 22440  
 ctgaagcaaa tctcagagt gcttcttcgg aggtttccaa ttcttggtt ccagcaccg 22500  
 ctacgcggte tggcgacca gactgacatt ggccccatg gtcgcgtga ccaatgaatt 22560  
 gatggcggt tggcgctcat cagcatccag gctagggtca ttctccagga tatcgatctc 22620  
 ccgttgagcg ccatccacat tattgccgat atcggattta gcttgcata tcaaccggc 22680  
 aatatgcctg tgccaggtaa tcaccgtggc gagataatc tgcagcgtca tcaattgatt 22740  
 gatgtttgca cccaggcgcc cgttggcagc attggcggg ccgcccggac ataggccg 22800  
 ttcgaagac tggcctttct gctggcgga ggtgtccaa acatcggtga cctttgcaa 22860  
 aacctggcta tattcctggg cccggtcata gaaagtgtt tcatcggtt ccaccagcc 22920  
 gcccggtacc agcatctgtc tggcatagct gcccgctggc ctggttaata tcatccccta 22980  
 ctgcccctcc caaacgccca gatcgctcgc cggatcacgc tccggttgg ctcgggatt 23040  
 tcacgcggc tgggcgctg gatccacccc gcgcgggtat tcgcagtaac ccgttgaatc 23100  
 cgcgcgcatg atgcaccgt tgggcagctc gccgggtgt cactcgctt gcgctggcg 23160  
 cgctgtcgca cggggcgctc ggtggtaac gacgtcataa ttaaccagc taaccgaacc 23220  
 taagaccagc tagctgcggc aatattggcg accaggacta tggcgccctc cgaaccggc 23280  
 cgatccatgt caaaacattg acaatgcgt ctacgcggt gtcggggcg ctgaatgacc 23340  
 gcattgcgg gctcattcgg tgcgtagtc ctaccaccg aacaatggg ttaggccatt 23400  
 ccttcggtca tcgcgcggga catggccgat aacgcagcg tcagctgctc gcccgcccg 23460

tcgttatacgc	cggacgcccgc	ggcctgcgca	ttgtgcagcg	cctcgttgac	ccgctgagcc	23520
accgcctcgg	caccagcctt	cttcagcaaa	ccatcttcga	tgcgagggcc	ggtgagccac	23580
ttgtgcccac	tgatcgtcac	ttcgacgggc	tcggcttcgt	cggtggcgcg	gaaggatccg	23640
ttgttcatct	gattgagcgt	cccgctctagg	gccgactgaa	accgcgcccgc	cagcgtcaac	23700
gccccgggcga	catgcgggtc	caattcgtcc	atgctcactt	cgactcctta	ctgtcctggc	23760
gccgacgggt	accaatgacg	gcctcgggtcc	atgcccgcac	ctcgggtgtg	agcgccctcgt	23820
cttcctgctg	agaacccttg	gacttggcgcc	ccccttggtcc	ctgatgcgcg	gcacccatcg	23880
gcattcccat	gccaccgccc	cccagcgcg	cgccgcgcgc	ggcccttccc	tggcctaagc	23940
cggcaatgtc	accagcgcca	gcggggcgca	cggattcggc	gccccgcac	gcggatccca	24000
acggcgccga	cggcaccgcc	ccgcctccac	cgccaccgag	cgatgcgcgt	ttgaccgcca	24060
cgtcgcccga	cagcgctgcg	gcttcccgcc	cagccgacgt	cagctgcgcg	gccgtgtcag	24120
ccgggaggcc	accaccggcg	gatccggtag	gcggaaccat	cggtgcgggt	ggcatcccg	24180
taccgggagt	cacaccggag	ccgtcagacg	gcggcatcag	gaagccagg	atcaatccct	24240
gctcttgccg	aggcgggggc	gggtcgatct	tgatggcggg	gggaggcttc	ggcggtttta	24300
ccggttccag	ggctgccttg	ttgttgattt	cggtcagcac	cttctccgac	ctctgctgat	24360
actccgcgta	caccgggaga	atgttggtcg	gggcccgaag	gttttccgcg	taaagccgtt	24420
cgagcccgcac	tatgtcttca	taagtcggat	gttcccgcct	agcccacacg	tgacgctgcg	24480
cgacatattg	agcctgcttg	gccatcgca	cgctcaattt	ggccatgtgg	agtatccatt	24540
gccgttggtg	atcgagcgaa	gcctcgcaag	cggtagccgc	atcgcccttc	cagttgtcaa	24600
acccccggaa	ccgcttgacg	tcgccttgca	gcgtcagggt	gaaagtgttc	cacccatccg	24660
caaagtgcgc	gagcgatgcg	ccttggtcgc	ccgtttcgag	cttccttgcc	gcttctttga	24720
gatccatgaa	gttgggttca	ccggccggtg	ccaccctcgg	cgtatcggtt	agttcgggcg	24780
aactgtcccc	tccgacggcc	ccggccgatt	ctgcctgcac	agttccttcg	ccgtcgttgt	24840
ccagcgcggt	cgcagcctcc	tcatacaact	cgccatacgc	cttgcccgcg	ttgcgcagcg	24900
aggtcgccag	acgtgcgcgc	tctttggcac	cgcccgccag	gtattcccgc	atgttgctcg	24960
cggacaatac	cagctggttg	gcggcggttt	tagccgcctg	gagttcgcac	ggtgtgatgg	25020
ggacatcagt	cggtgggtcc	gccatcgggg	cctccacctc	gttgggccctg	ttcaaaatct	25080
cttgctgata	caccgtcacg	gtctgcgact	gcctcatatc	ggatcatcct	ccttagtgct	25140
atagccatta	tcgtcgctaa	actgaaaggt	tcctgcacta	atttgatgcc	gcccgttcat	25200
gccggcactc	cgaacggatc	gccctacttc	ggcagcgcca	tctggtagcg	gctttcctcg	25260
ggtggggaaa	cccgggcaat	cgccagctgc	cgatgcgcgc	gggtaccgat	cacattgtgc	25320
cgcagaatca	cccgggtcaat	accgggatgc	gggcccagat	aggctcgtcg	attcgggcac	25380
gccaccttta	cctcctgccc	gatgtgtgcg	ccgatcaacc	gggcaaattc	ctcgaactgt	25440
ggcccgcactg	tgaccatcgc	acctgccgcg	gccgcacgca	ccacgaactg	ggtgaatgtc	25500
tgagcgtcac	ccaggttgag	ggcgatgtcg	acatcgtcga	agggcatgta	gaccgggcat	25560
cggttcaccg	tctcgccgac	cagtaaccca	gctgaccgca	tcggcagctg	gcagtggcg	25620
ttggccacca	gatgtgggcc	ttgcagcgcg	ggcgctgcgc	cgccaaatag	gcggggcgag	25680
ccccctgggtg	tcttgggctt	gtccgcgctg	gtcagcaaca	ccgtggactg	cgggggccatc	25740
cccggcgcg	cccggactct	ggtgatgggt	tggtccgcgc	gcgcgcacca	ccatacatcc	25800
ggacctccgg	gcgcgcgcta	ggcggcagtg	taggcacgcg	gccccttgat	catcgaccat	25860
ttctcccgca	caaagccgat	gtcgggtggc	tggtcgtagt	catcgaagct	gcggccacac	25920
accgcgtcga	caccatggct	agccagtcga	tcggcaatgc	gcgtcgcgga	cgccaccaaaa	25980
taccgggcca	gtcctgcgac	gccttcacgc	cggcgctgcg	ccgatttgcg	ggtgcgttcc	26040
gggtcggcg	gcagcacgat	ccaggtccgg	cggttcgcgc	gcgcgcgggtc	tgtcccgcac	26100
acctgctgat	acagactcac	cacgtccggc	gctgcgggtat	tgccgacgcg	gtagccgggt	26160
gagacgatat	cggcctccaa	gtcgggacag	tgacccgaca	ggagctcctc	caccagtcgg	26220
gtgtccagca	tgctcgtcgt	gtgggcttgc	ccgtcgacga	tgaccgtcgg	cgtgaatggt	26280
cggggaatga	gtcgcattac	ggcgaccaga	aactcgctt	gccagcgcac	cgcaacgtga	26340
tctcctgggt	tcacgggtgg	cccgaaccac	ggttctgacg	aggaatccgg	gggcccgtcg	26400
cgcgcgcgca	accacgcgta	caccgcgcgc	acccagccgg	tgatccggcg	gccgtagaaa	26460
gtgaccgtgg	ccacgatgac	gcccacagag	gccagcgcaa	tcccgcccca	ccagtagcgc	26520
gtctccaaga	atgcgatgat	gcattggcg	gccaacgcgg	aggcaagcaa	ggcgtgccc	26580
gtgctgaacc	gcagccctaa	aggattttct	atcgccgggt	cagcgcccgt	ctagccagcg	26640
cgcgccaggcc	cagggccaac	gtaaggccga	cgggccacca	cgccacagcc	gtaatcgggc	26700
gacgatcggg	acccggctcc	accaccgggg	gtgggaagtc	tctgacgttg	tatggcgccg	26760
aagcaggggc	gggcggaatg	tcccacgtca	gcgcggccac	cgcacgatg	acgcggcg	26820
cgaccaggct	gtcgaccccg	ccccgggggt	gtctcgcggt	ggcggtgatc	cggtggatga	26880
tctgcgcggg	cgctcaggtcg	gggaaccgct	gccgaagcag	ggccgcgaga	cccgcacacat	26940
atgccgcggc	aaacgaggtg	ccggcgatgg	gtaccggccc	ctcccgccct	tgacgcgcac	27000
tcaccgggtt	accgggtgct	ccgagcgcg	cgatgttttc	tgcgggcgcg	gccacgtcca	27060
cccacgggtc	gtgcatcgag	aacgagctgg	gcaccccggt	ctggccgata	cggccgacgc	27120



ttaacaccag	cgggtgcgtac	cacgccgggg	tgacaacgggt	ctgcacattg	ttccagccgc	27180
gtgggtcgcc	gggtgtggac	gggtccggcg	cgggattctg	tacgcaatcg	ccaccggtgt	27240
tgccggccgc	gaccaccacc	accacgcctt	tgacgttgac	cgcatagtcg	atggatgcac	27300
ccagtggagt	ttcatcgatc	ggcctgctca	ccttgttagca	ggcggcttca	ctgatgttga	27360
tcacacccac	gccgaggttg	gcggcggtgca	ccacggcgcg	ggcaagactg	cggatggaac	27420
cggcgggccg	gggtggcggtt	gggtcattcg	ggttggcctt	tgagccgacc	ggttcgaagg	27480
cctcagacgt	ctgacgtagc	gagagcagtc	gagcgctcggg	cgcgacgcgc	acgaacccgt	27540
cgggtgggcgc	ggggccggccc	gcgatgatgg	atgctgtgag	agtcccatgg	gcatacacagt	27600
cagacaggcc	gttaccggcc	tggtcgacga	aatcgccgcc	aggttccgcc	gggacccgtg	27660
gcgaagcgtc	gacaccgggtg	tcgatcaccg	ccacgcgtcac	cccggccccc	gtcgcgaact	27720
tgtgggcatc	ggccacgcgc	agatacgtgt	tgctccacgg	cggatcgtag	aaccgcgacc	27780
ccggcagcgt	gggtgggcgac	gcgcacaaaa	cgcgctgttc	ggtaggctga	tccgggcccc	27840
tcacgtcggg	cggcaacgcg	cccggatcga	tcggcggttg	cgtgatggcc	gatgcgggcg	27900
acgcggtgag	caacgccagc	gccaccgtga	tcagaaagat	acggtgcact	cccagaacac	27960
tccattcggt	gagattcatt	gcgattcatt	gagctgcgtt	gctaccttgg	gccacttgac	28020
ggacctgtgt	gcatttttaga	cgtaacggct	gggcaaacaa	cgtgttcacg	cctgggctgg	28080
tccgcgcgcg	cgaccagggc	gcgtaggcgc	tgtacctgga	ccacgccggg	actcaacggt	28140
tttgctaccg	cactagccga	tatgcggctg	ctaccaaacc	atcgccggcca	tgtctcggtt	28200
gtctgagcac	acgctgcgta	tcgcggcatc	gatgtcggtg	gcggtgatga	tctgcagatc	28260
ctgaaccgat	accggttggc	ccgcacggtt	tgcgcgaacc	accggggtgt	cccgggaacc	28320
ttcggcgcg	tcgatcacgt	tgccggcgaa	ccgacggctt	tgcatagcgt	cgataccgtg	28380
ctgcccacta	gggggtggtg	agttacggat	gggtggtgacc	gcgtcgagga	atacctcccg	28440
tgccgcgtca	tcgagctggc	tggcgcgcg	tgtagcgtag	cggtgtccaa	tctcgacgat	28500
ctccaccggc	gaataagact	cgaaccgcag	ctttcggttg	aaccggccag	ccaaacccgg	28560
gttcacgggtg	aggaattcat	ccacctgac	ctcatagccg	gccccgatga	aacagaagtc	28620
gaatcggtgt	gtttccaatt	gaaccaggag	ttgattgacc	gcctccatgc	cgatcatgtc	28680
cgggtgttccg	tcttgatgac	gttcgatcag	cgagtagaac	tcgtccatga	aaatgattcg	28740
cccgagtgc	ttttcgatca	gctcggtcgt	cttgggtcct	gactccccga	tgtagtgcgc	28800
acagaagtcc	gatcggcgaa	cttctcgaat	tctcgggtga	cgcacgatcc	ccatgcggcg	28860
gtagatcttg	ccgagcgctt	tgctttacct	tgctttacct	gtgcctgggtg	gccccaccag	28920
caacatgttg	ttggtctgcc	cctccaccgg	taggccgtgc	tctaggcgca	tcatgcgcac	28980
ctcgagttgg	tcttccagcg	ccgataccgc	ttgcttgacc	gccgccaggc	ccacctgttt	29040
ggccagcagt	tccccggcct	cggctagcag	ctcgccgcgc	cgtgcgctg	cattgtcgtc	29100
atcgagctgg	tcgcggcctt	tcgccgtcga	agcatcccaa	cggtcggagc	ggctggcgat	29160
ggttcgttca	tcggtaacaa	tcaagcgag	gttcgggtcc	gccagggctt	ctttggcggc	29220
gtcggtgagc	accccggttg	tggtggcctt	cgacagccag	atctgggcct	tgtcctcctc	29280
atgcagttgc	cggtaacacca	tcccccgcac	atacgccaag	tcggcgacca	gcagcggaat	29340
atcgcccggt	ccgatcgccg	cggtagcacc	gtcggcgccg	aaccgctccg	atgacctgct	29400
gtgtccgatc	acgtccaccc	ggtcacgcca	gtccaggggc	actcgcccct	gcccagatg	29460
ggcggcgccg	tgggctgcca	gcgcacaaat	cgacgcgggtc	accgccggca	tgacgatcgc	29520
ctgtggcgcc	agatcctcgg	cggccgtcga	caacacgtcg	ggccatcgct	gcgtgacgta	29580
catcaggaac	gcccagagcca	gctgatgcca	ctggtagttg	cgccacgaat	ccaatagctc	29640
gcggtttgct	aacagggcat	cggccttcgc	atactcccc	gcgatcgtea	acgccgacga	29700
cagcgccagc	cccacctgag	atgcgtcggt	caccgtgatc	ccgatggatg	gtcccagctg	29760
gacctcagcg	gccaacgtcc	ggccgatccg	cgtgggtctcg	cgggtgcagcc	actcgctatg	29820
ggcggttagc	tgcttaagcg	aggccagatc	gcggtcaccg	caggcgatac	gacccagcca	29880
cgcgctcgcc	atcgacggat	cggcctcggt	ggcagccaca	aactcaggca	acgcgcggcc	29940
gcaccccttg	ccattcttga	tcgtcatcgc	ccgatcgaaa	tgccggcgcg	cagttagtaa	30000
atcacccatc	gtgtccacca	ttctcgacat	cgccgcgcgt	gtcaccgcgg	ttgcaacgtg	30060
tgtctgtcac	tctgtgcctc	aaattccggt	ggcaacgttc	taccggccta	tcgacatcgt	30120
gaccggctca	aggctgacat	agcggttctc	cgcacggaac	atttccatct	caaccagcca	30180
gttttgtcct	gccgcaccga	ctttaccggt	tgcccgatcg	atttgttcga	tggtcacctc	30240
gaagccatgc	cgatcgctct	cggacagcga	ggtaccgggt	cgggcaatgg	tgatgacact	30300
ggctggccgt	ggcgtgggcg	aaatcgcgac	atcgacaccg	ctgccttcag	atttgccgtc	30360
atcgccgttc	ttgcgcccgc	gcacgtactc	cacgacgcgc	acagtgggtg	gcggcgcggg	30420
cctgtggtg	ccgacgatgc	tcaactgcgg	catcgctacg	ctggcccaac	gctcttggtc	30480
gcgagtgtgc	acacacaccc	gctcacccgg	accgacgacg	cgaatcacga	tcctcttggtc	30540
gatcggtgcg	tccgcggcca	cgaagacgcg	cgacagctca	ccggcgctcg	taacgggaat	30600
catcagccgg	tccccgttgc	tcagcttgcc	aatcaacacc	cccgcggtgc	cgatctcggt	30660
gactagctgc	gccggcaacg	ggcagcgccg	ctgtccgcgt	aggtgtggac	gtggcccgcga	30720
catggtggcc	gcagccgcgg	cggcttgctc	accattgagc	cgacgcaaga	tcacactggg	30780

```

cggggtaggc gccggcgctcg gtgtgcgcac ggtgatggtc gcggtgcacg tcgctgccgg 30840
atacaccgtt acgttcttga tgacctcatc ggcacgcagc gtccaggctt gcgagagaac 30900
ccgcgacgaa atcgccctcag ccgggtacgc atacgtcgtc atccaccggg cttcaccggc 30960
gatagctttc cagcgctgcg cactcccggc taccgcgtcc gaccccagcc ggcgatcaag 31020
ctcagccaag tctgttgccg tggccagttt ggcgcgcaag ccctgacagc gcagggagct 31080
ggcaacgcgt tgggcgaccg aaatggcagc ggccccaacg ctggtacgcc agcgtaaagc 31140
ttgggtgttg ccgatcaccg gaagccgcat gatcagccac gtttcgcgcc gcccggcata 31200
cgccggcgta ccgatctccg cgtcatacac ccgcgggtaa tcgccgacgg tgccggttcg 31260
cgagccgaag gtgacgacgc tgattgaatc gaggttccagg tccagcgggt ggcgagcaa 31320
cggcgcgagc tcaacgacgt caatcacgtt gtcgctttct acggtcaccg acccggtgac 31380
cgtagtcgcc cgggtgcgtc ggccgagaag ttgcaccgcc accaccgcga caccgtcttg 31440
cacgcggacg ccacccccgg atcgggtgtt ggccaaggta attgggtcat tccatttgac 31500
gggacggcga ccccgacgcc ccagtagcgc ccacgaccac gccggctgac cccaccactg 31560
tacgaacacc aaggcgacgc cgaccacgac agccatgacc gcacctagct ggccgcccag 31620
cgcccagccc gccgacgcga gcacgaacac tgtccacacc ccggcgaccc gcctcgact 31680
gcgcgggctg aacccgggtc gcttggacgt caacgcgccc tccgtagccg agccccgatt 31740
gccattgcca gcacaccggg ggccactgcg ccgacgaacc cgatagcgat attgcgcgcc 31800
cggtgac 31808

```

<210> 2

<211> 13773

<212> DNA

<213> mycobacterium tuberculosis

<220>

<223> Complete DNA sequence of RD1 Rv3867-3877

<400> 2

```

atgggtggacc cgccgggcaa cgacgacgac caggtgatc tcgacgccct cgatttctcc 60
gccgccacaca ccaacgaggg gtgcgcgctg gacgccttag acgactatgc gccggtgcag 120
accgatgacg ccgaaggcga cctggacgcc ctccatgcgc tcaccgaacg cgacgaggag 180
ccggagctgg agttgttcac ggtgaccaac cctcaagggt cgggtgcggt ctcaaccctg 240
atggacggca gaatccagca cgtcgagctg acggacaagg cgaccagcat gtccgaagcg 300
cagctggccg acgagatctt cgttattgcc gatctggccc gccaaaaggc gcgggcgtcg 360
cagtacacgt tcatggtgga gaacatcggg gaactgaccg acgaagacgc agaaggcagc 420
gccctgctgc ggggaattcg ggggatgacc ctgaatctgc cgacgccgga agaggctgcc 480
gcagccgaag ccgaagtgtt cgccaccgcg tacgatgtcg actacacctc ccggtacaag 540
gccgatgact gatcgcttgg ccagtctgtt cgaagcgccc gtcagcatgt tgccgatgtc 600
ggaggcgccg tcgctagatc tgttcaccga gatcaccaac tacgacgaat ccgcttgcca 660
cgcatggatc ggccggatcc ggtgtgggga caccgaccgg gtgacgctgt ttgcgcctg 720
gtattcgccg cgcaatttcg gacagtgtgc gggatcggtc cagatctcga tgagcacgtt 780
aaacgccagg attgccatcg gggggctgta cggcgatatc acctaccggg tcacctcgcc 840
gctagcgatc accatgggct ttgccgcatg cgaggcagcg caaggcaatt acgcgcagcg 900
catggaggcc ttagaggccg ccccggtcgc gggttccgag cacctggtgg cgtggatgaa 960
ggcggttgtc tacggcgccg ccgaacgctg gaccgacgtg atcgaccagg tcaagagtgc 1020
tgggaaatgg ccggacaagt ttttgccgg cgccggccgt gtggcgcacg gggttgccgc 1080
ggcaaacctg gccttggtca ccgaagccga acgcccactc accgaggcca acgactcgcc 1140
cgccgggtgag gcgtgtgcgc gcccatcgc ctggtatctg gcgatggcac ggccgagcca 1200
gggcaacgaa agcgcccgcg tggcgctgct ggaatggtta cagaccactc accccgagcc 1260
caaagtggct gcggcgctga aggatccctc ctaccggctg aagacgacca ccgccgaaca 1320
gatcgcatcc cgcgccgatc cctgggatcc gggcagtgct gtgaccgaca actccggccg 1380
ggagcggctg ctgcgccagg cccaagccga actcgaccgc caaattgggc tccccgggt 1440
taaaaatcag attgaacgct accgcgcggc gacgctgatg gcccgggtcc gcgccgcaa 1500
gggtatgaag gtcgcccagc ccagcaagca catgatcttc accggaccgc ccggtaccgg 1560
caagaccacg atcgcgccg tgggtggcaa tatctggcc ggttagggc tcatgcccga 1620
acccaactc gtcgagcgt cgcgcaagg cttcgtcgcc gagtacgagg ggcaatcgcc 1680
ggtcaagacc gctaagacga tcgatcaggc gctgggcggg gtgcttttca tcgacgaggc 1740
ttatgcgctg gtgcaggaaa gagacggccg caccgatccg ttcggtaag aggcgctgga 1800
cacgtgctg gcgcggatg agaacgaccg ggaccggctg gtggtgatca tcgccgggta 1860
cagctccgac atagatcggc tgctggaaac caacgagggt ctgcggtcgc ggttcgccac 1920

```

tcgcatcgag	ttcgacacct	attcccccca	ggaactcctc	gagatcgcca	acgtcattgc	1980
cgctgctgat	gattcggcgt	tgaccgcaga	ggcgcccgag	aactttcttc	aggccgcca	2040
gcagttggag	cagcgcatgt	tgccgcccgt	gcgcgcccgt	gacgtcgccg	gcaacggtcg	2100
gtatgcgcgc	cagctgggtg	aggccagcga	gcaatgccgg	gacatgcgtc	tagcccaggt	2160
cctcgatata	gacaccctcg	acgaagaccg	gcttcgcgag	atcaacggct	cagatatggc	2220
ggaggctatc	gccgcgggtg	acgcacacct	caacatgaga	gaatgaacta	tggggcttcg	2280
cctcaccacc	aaggttcagg	ttagcggctg	gcgtttctcg	ctgcgcccgg	tcgaacacgc	2340
catcgtgcgc	cgggacaccc	ggatgtttga	cgaccccgct	cagttctaca	gccgctcgat	2400
cgctcttggc	atcgtcgtcg	cggtcctgat	tctggcgggt	gccgcgctgc	tggcgtactt	2460
caaaccacaa	ggcaaactcg	gcggcaccag	cctgttcacc	gaccgcgcga	ccaaccagct	2520
ttacgtgctg	ctgtccggac	agttgcaccc	ggctctacaac	ctgacttcgg	cgcggtgggt	2580
gctgggcaat	ccggccaacc	cggccaccgt	gaagtcctcc	gaactgagca	agctgccgat	2640
gggcagacc	gttgggaatc	ccggcgcccc	ctacgccacg	cctgtttcgg	cgggcagcac	2700
ctcgatctgg	accctatgcg	acaccgctcg	ccgagccgac	tccacttccc	cggtagtgc	2760
gaccgcggtc	atcgcgatgc	cgttggagat	cgatgcttcg	atcgatccgc	tccagtcaca	2820
cgaagcgggt	ctggtgtcct	accagggcga	aacctggatc	gtcacaacta	agggacgcca	2880
cgccatagat	ctgaccgacc	gcgcctccac	ctcgtcgatg	gggataccgg	tgacggccag	2940
gccaaacccc	atctcggagg	gcatgttcaa	cgcgctgcct	gatatggggc	cctggcagct	3000
gccgcggata	ccggcgggcg	gcgcgcccac	ttcgcttggc	ctacctgatg	atctagtgat	3060
cggatcggtc	ttccagatcc	acaccgacaa	gggcccgcga	tactatgtgg	tgctgcccga	3120
cggcatcgcg	caggtcaacg	cgacaaccgc	tgcggcgctg	cgcgccaccc	aggcgacagg	3180
gctggtcgcg	ccaccggcaa	tggtgcccag	tgtgtcgtc	agaatcgccg	aacgggtata	3240
ccctcaccg	ctaccggatg	aaccgctcaa	gatcgtgtcc	cgcccgccag	atcccgcgct	3300
gtgctggtca	tggcaacgca	gcgcggcgga	ccagtcgccg	cagtcaacgg	tgctgtccgg	3360
ccggcatctg	ccgatatacg	cctcagcgat	gaacatgggg	atcaagcaga	tccacgggac	3420
ggcgaccggt	tacctcgacg	gcggaaaatt	cgtaggcactg	caatcccccg	atcctcgata	3480
caccgaatcg	atgtactaca	tcgatccaca	ggcgctgcgt	tatgggggtgc	ctaacgcgga	3540
gacagccaag	tcgctggggc	tgagttcacc	ccaaaacgcg	ccctggggaga	tcgttcgtct	3600
cctggtcgac	ggtccgggtg	tgctcgaaaga	tgccgcactg	ctcgagcacg	acacgctgcc	3660
cgctgaccct	agcccccgaa	aagttcccgc	cggagcctcc	ggagccccct	gatgacgacc	3720
aagaagttca	ctcccaccat	taccctgtgg	ccccggttga	ccccggcgga	gatcagcctc	3780
acgcgcggcg	atgacctggg	catcgacatc	ccaccgctcg	gcgtccaaaa	gatccttccc	3840
tacgtgatgg	gtggcgccat	gctcggcatg	atcgccatca	tggtggccgg	cgccaccagg	3900
cagctgtcgc	cgtacatggt	gatgatgccg	ctgatgatga	tcgtgatgat	ggtcggcggt	3960
ctggccggta	gcaccgggtg	tgggcggaag	aaggtgcccg	aaatcaacgc	cgaccgcaag	4020
gagtagctgc	ggtatttggc	aggactacgc	acccgagtg	cgctcctcgg	cacctctcag	4080
gtggcggttc	tctcctacca	cgcaccgcat	cccgaggatc	tgttgctgat	cgtcggcacc	4140
caacggcagt	ggtcccggcc	ggccaacgcc	gacttctatg	cggccaccgc	aatcggtatc	4200
ggtgaccagc	cggcggttga	tcgattattg	aagccggccc	tcggcgggga	gttggccgcc	4260
gccagcgag	cacctcagcc	gttcctggag	ccggtcagtc	atatgtgggt	ggtcaagttt	4320
ctacgaaccc	atggattgat	ccatgactgc	ccgaaactgc	tgcaactccg	tacctttccg	4380
actatcgcca	tcggcgggga	cttggcgggg	gcagccggcc	tgatgacggc	gatgatctgt	4440
cacctagccg	tgttccaccc	accggacctg	ctgcagatcc	gggtgctcac	cgaggaaccc	4500
gacgaccccg	actggtcctg	gctcaaatgg	cttcgcgcag	tacagcacca	gaccgaaacc	4560
gatgcggccg	ggtccaccgc	gctgatcttc	acgcgccagg	aaggtctgtc	ggacctggcc	4620
gcgcgcgggg	cacacgcacc	cgattcgttt	ccggcgggcc	cctacgtagt	cgctcgtcgac	4680
ctgaccggcg	gcaaggctgg	attcccgcgc	gacggtaggg	ccggtgtcac	ggtgatcacg	4740
ttgggcaacc	atcgcggtc	ggcctaccgc	atcagggtgc	acgaggatgg	gacggctgat	4800
gaccggctcc	ctaaccaatc	gtttcgccag	gtgacctcgg	tcaccgatcg	gatgtcgccg	4860
cagcaagcca	gccgtatcgc	gcgaaagtgg	gccggatggg	ccatcacggg	caccatcctc	4920
gacaagacgt	cgcgggtcca	gaagaagggt	gccaccgact	ggcaccagct	ggtcgggtgc	4980
caaagtgtcg	aggagataac	accttcccgc	tgaggatgtg	acaccgacac	cgaccgtgac	5040
cggctaaaga	tcccgtttgg	tcatgaacta	aagaccggca	acgtcatgta	cctggacatc	5100
aaagagggcg	cggaattcgg	cgccggaccg	cacggcatgc	tcacggggac	cacgggggtc	5160
gggaagtccg	aattcctgcg	caccctgatc	ctgtcgtctg	tggcaatgac	tcattccagat	5220
caggtgaatc	tcctgctcac	cgacttcaaa	ggtggttcaa	ccttcctggg	aatggaaaag	5280
cttccgcaca	ctgccgctgt	cgtaaccaac	atggccgagg	aagccgagct	cgtcagccgg	5340
atgggcgagg	tgttgaccgg	agaactcgat	cggcgccagt	cgatcctccg	acaggccggg	5400
atgaaagtgc	gcgcggccgg	agccctgtcc	ggcggtggcc	aatacgagaa	gtaccgcgaa	5460
cgcggtgccc	acctaccccc	gctgccaaac	cttttcgtcg	tcgtcgacga	gttcgcccag	5520
ctgttgacga	gtcaccgcga	cttcacgggg	ctgttcgacc	ggatctgcgc	cgtcgggggg	5580

tcgctgaggg	tccatctgct	gctggctacc	cagtcgctgc	agaccggcgg	tgttcgcatc	5640
gacaaactgg	agccaaacct	gacatatcga	atcgattgc	gcaccaccag	ctctcatgaa	5700
tccaaggcgg	taatcggcac	accggaggcg	cagtatcatca	ccaacaagga	gagcgggtgc	5760
gggtttctcc	gggtcggcat	ggaagaccgc	gtcaagttca	gcaccttcta	catcagtggg	5820
ccatacatgc	cgccggcggc	aggcgtcgaa	accaatggtg	aagccggagg	gcccgggtcaa	5880
cagaccacta	gacaagccgc	gcgcattcac	aggttcaccg	cggcaccggt	tctcgaggag	5940
gcgccgacac	cgtgaccgcg	gccggcgacg	atgcaaagcg	cagcgatgag	gaggagcggc	6000
gcc'aacggcc	cgcgccggcg	acgatgcaaa	gcgcagcgat	gaggaggagc	ggcgcgcatg	6060
actgctgaac	cggaagtacg	gacgctgcgc	gaggttggtg	tggaccagct	cggcactgct	6120
gaatcgcggtg	cgtacaagat	gtgggtgccg	ccgttgacca	atccgggtccc	gctcaacgag	6180
ctcatcgccc	gtgatcgggc	acaacccctg	cgatttgccc	tggggatcat	ggatgaaccg	6240
cgccgccatc	tacaggatgt	gtggggcgta	gacgtttccg	gggcccggcg	caacatcggt	6300
attggggggc	cacctcaaac	cggaagtcg	acgctactgc	agacgatggt	gatgtcggcc	6360
gccgccacac	actcaccgcg	caacgttcag	ttctattgca	tcgacctagg	tggcggcggg	6420
ctgatctatc	tcgaaaacct	tccacacgtc	ggtggggtag	ccaatcggtc	cgagcccgcac	6480
aaggtcaacc	gggtggtcgc	agagatgcaa	gccgtcatgc	ggcaacggga	aaccaccttc	6540
aaggaacacc	gagtgggctc	gatcgggatg	taccggcagc	tgcgtgacga	tccaagtcaa	6600
cccgttgctg	ccgatccata	cggcgacgtc	ttctgatca	tcgacggatg	gcccggtttt	6660
gtcggcgagt	tccccgacct	tgagggggcag	gttcaagatc	tggccgccc	ggggctggcg	6720
ttcggcgctc	acgtcatcat	ctccacgcca	cgctggacag	agctgaagtc	gcgtgttcgc	6780
gactacctcg	gcaccaagat	cgagttccgg	cttgggtgacg	tcaatgaaac	ccagatcgac	6840
cggattacc	gcgagatccc	ggcgaatcgt	ccgggtcggg	cagtgtcgat	ggaaaagcac	6900
catctgatga	tcggcggtgc	caggttcgac	ggcgtgcaca	gcgccgataa	cctggtggag	6960
gcgatcacgc	cgggggtgac	gcagatcgct	tcccagcaca	ccgaacaggc	acctccggtg	7020
cggttcctgc	cggagcgtat	ccacctgcac	gaactcgacc	cgaacccgcc	gggaccagag	7080
tccgactacc	gcactcgctg	ggagattccg	atcggttcgc	gcgagacgga	cctgacgccg	7140
gctcactgcc	acatgcacac	gaacccgcac	ctactgatct	tcggtgcggc	caaatcgggc	7200
aagacgacca	ttgcccacgc	gatcgcgcg	gccattttgtg	cccgaacag	tccccagcag	7260
gtgcggttca	tgctcgcgga	ctaccgctcg	ggcctgctgg	acgcggtgcc	ggacacccat	7320
ctgctgggcg	ccggcgcgat	caaccgcaac	agcgcgtcgc	tagacgaggc	cgttcaagca	7380
ctggcggtca	acctgaagaa	gcggttcg	ccgacgcacc	tgacgacggc	gcagctacgc	7440
tcgcttctcg	gggtggagcg	atttgacgtc	gtgcttcggtg	tcgacgattg	gcacatgac	7500
gtgggtgccg	ccggggggat	gccgcgatg	gcaccgctgg	ccccgttatt	gccggcgggc	7560
gcagatatcg	ggttgcacat	cattgtcacc	tgtcagatga	gccaggctta	caaggcaacc	7620
atggacaagt	tcgtcggcgc	cgcatcggg	tcgggcgctc	cgacaatgtt	cctttcgggc	7680
gagaagcagg	aattcccatc	cagtgaagtc	aaggccaagc	ggcgcccccc	tggccaggca	7740
tttctcgtct	cgccagacgg	caaagaggtc	atccaggccc	cctacatcga	gcctccagaa	7800
gaagtgttcg	cagcaccccc	aagcgccggt	taagattatt	tcattgccgg	tgtagcagga	7860
cccagctca	gcccggtaat	cgagttcggg	caatgctgac	catcgggttt	gtttccggct	7920
ataaccgaac	ggtttgtgta	cgggatacaa	atacaggag	ggaagaagta	ggcaaatgga	7980
aaaaatgtca	catgatccga	tcgctgccga	cattggcagc	caagtgaagc	acaacgctct	8040
gcacggcggtg	acggccggct	cgacggcgct	gacgtcgggtg	accgggctgg	ttcccgcggg	8100
ggccgatgag	gtctccgccc	aagcggcgac	ggcgttcaca	tcggagggca	tccaattgct	8160
ggcttccaat	gcacggccc	aagaccagct	ccaccgtgcg	ggcgaagcgg	tccaggacgt	8220
cgcccgacac	tattcgcaaa	tcgacgacgg	cgccgcggcg	gtcttcgccc	aataggcccc	8280
caacacatcg	gagggagtg	tcaccatgct	gtggcacgca	atgccaccgg	agctaaatac	8340
cgcacggctg	atggccggcg	cggttccggc	tccaatgctt	gcggcgggcg	cgggatggca	8400
gacgttttcg	gcggtctctg	acgctcaggc	cgtcgagttg	accgcgcgcc	tgaactctct	8460
gggagaagcc	tggactggag	gtggcagcga	caaggcgctt	gcggctgcaa	cgccgatggt	8520
ggtctggcta	caaaccgcgt	caacacaggc	caagaccgct	gcgatgcagg	cgacggcgca	8580
agccgcggca	tacaccaggg	ccatggccac	gacgccgtcg	ctgccggaga	tcgccgcca	8640
ccacatcacc	caggccgtcc	ttacggccac	caacttcttc	ggtatcaaca	cgatcccgat	8700
cgcgttgacc	gagatggatt	atttcacccg	tatgtggaac	caggcagccc	tggcaatgga	8760
ggtctaccag	gccgagaccg	cgtttaacac	gcttttcgag	aagctcgagc	cgatggcgct	8820
gatecttgat	cccggcgcg	gccagagcac	gacgaacccg	atcttcggaa	tgccctcccc	8880
tggcagctca	acaccggttg	gccagttgcc	gccggcggtt	accagacccc	tcggccaact	8940
gggtgagatg	agcggcccg	tgacgagct	gaacccagcg	ctgcagcagg	tgacgtcggt	9000
ggtcagccag	gtggcgggca	ccggcgcg	caacccagcc	gacgagggaag	ccgcgcagat	9060
gggcctgctc	ggcaccagtc	cgctgtcgaa	ccatccgctg	gctggtggat	caggccccag	9120
cgcgggcgcg	ggcctgctgc	gcgcggagtc	gctacctggc	gcaggtgggt	cgttgacccg	9180
cacgccgctg	atgtctcagc	tgatcgaaaa	gccgggttgc	ccctcggtga	tgcggcgggc	9240

tgctgcggga	tcgtcggcga	cggtgggcgc	cgctccgggtg	ggtgcgggag	cgatggggcca	9300
gggtgcgcaa	tccggcggct	ccaccaggcc	gggtctgggtc	gcgcccggcac	cgctcgcgca	9360
ggagcgtgaa	gaagacgacg	aggacgactg	ggcgaagag	gacgactggt	gagctcccgt	9420
aatgacaaca	gacttcccgg	ccaccggggc	cggaagactt	gccaacattt	tggcgaggaa	9480
ggtaaagaga	gaaagtagtc	cagcatggca	gagatgaaga	ccgatgccgc	taccctcgcg	9540
caggaggcag	gtaatttcga	gcggatctcc	ggcgacctga	aaaccagat	cgaccagggtg	9600
gagtcgacgg	caggtttcgtt	gcagggccag	tggcgcgccg	cggcggggac	ggccgcccag	9660
gccgcggtgg	tgcgcttcca	agaagcagcc	aataagcaga	agcaggaact	cgacgagatc	9720
tcgacgaata	ttcgtcaggg	cggcgtccaa	tactcgaggg	ccgacgagga	gcagcagcag	9780
gcgctgtcct	cgcaaatggg	cttctgaccc	gctaatacga	aaagaaacgg	agcaaaaaca	9840
tgacagagca	gcagtggaa	ttcgcgggta	tcgaggccgc	ggcaagcgca	atccagggaa	9900
atgtcacgtc	cattcattcc	ctccttgacg	aggggaagca	gtccctgacc	aagctcgcag	9960
cggcctgggg	cgttagcggt	tcggaggcgt	accagggtgt	ccagcaaaaa	tgggacgcca	10020
cggtaccga	gctgaacaac	gcgctgcaga	acctggcgcg	gacgatcagc	gaagccgggtc	10080
aggcaatggc	ttcgaccgaa	ggcaacgtca	ctgggatgtt	cgcatagggc	aacgccgagt	10140
tcgcgtagaa	tagcgaaaca	cgggatcggg	cgagttcgac	cttcctcgcg	tctcgccctt	10200
tctcgtgttt	atacgtttga	gcgcactctg	agaggttgct	atggcgcccg	actacgacaa	10260
gctcttccgg	ccgcacgaag	gtatggaagc	tccggacgat	atggcagcgc	agccgttctt	10320
cgaccccagt	gcttcgtttc	cgccggcgcc	cgcatcgcca	aacctaccga	agcccaacgg	10380
ccgactccg	cccccgacgt	ccgacgacct	gtcggagcgg	ttcgtgtcgg	ccccgcggcc	10440
gccaccccca	cccccacctc	cgctctcgcc	aactccgatg	ccgatcgccg	caggagagcc	10500
gcctcgcgg	gaaccggcgg	catctaaacc	accacacacc	cccattgccc	tcgcccggacc	10560
cgaaccggcc	ccaccgaaac	caccacacacc	ccccatgccc	atcgccggac	ccgaaccggc	10620
cccacccaaa	ccaccacac	ctccgatgcc	catcgccgga	cctgcacca	ccccaaccca	10680
atcccagttg	gcgcccccca	gaccaccgac	accacaaacg	ccaaccggag	cgccgcagca	10740
accggaatca	ccggcgcccc	acgtaccctc	gcacggggca	catcaacccc	ggcgacccgc	10800
accagcaccg	ccctgggcaa	agatgccaat	cggcgaaccc	ccgcccgcctc	cgctccagacc	10860
gtctgcgtcc	ccggccgaac	caccgacccg	gcctgcccc	caacactccc	gacgtgcgcg	10920
ccggggtcac	cgctatcgca	cagacaccga	acgaaacgtc	gggaaggtag	caactgggtc	10980
atccatccag	gcgcggctgc	gggcagagga	agcatccggc	gcgcagctcg	cccccggaac	11040
ggagccctcg	ccagcgccgt	tgggccaacc	gagatcgat	ctggctccgc	ccaccggccc	11100
cgcgccgaca	gaacctcccc	ccagccccctc	gcgcgacgcg	aactccgggtc	ggcgtgccga	11160
gcgacgcgtc	caccccgatt	tagccgcccc	acatgccgcg	gcgcaacctg	attcaattac	11220
ggccgcaacc	actggcggtc	gtcgccgcaa	gcgtgcagcg	ccggatctcg	acgcgacaca	11280
gaaatcctta	aggccggcgg	ccaagggggc	gaaggtgaag	aaggtgaagc	cccagaaacc	11340
gaaggccacg	aagccgcccc	aagtgggtgtc	gcagcgcggc	tggcgacatt	gggtgcatgc	11400
gttgacgcga	atcaacctgg	gcctgtcacc	cgacgagaag	tacgagctgg	acctgcacgc	11460
tcgagtccgc	cgcaatcccc	gcgggtcgta	tcagatcgcc	gtcgtcggtc	tcaaagggtg	11520
ggctggcaaa	accacgtcga	cagcagcggt	ggggtcgacg	ttggctcagg	tgcggggcca	11580
ccggtacctg	gctctagacg	cggatccagg	cgccggaaac	ctcgccgatc	gggtaggggc	11640
acaatcgggc	gcgaccatcg	ctgatgtgct	tgcagaaaaa	gagctgtcgc	actacaacga	11700
catccgcgca	cacactagcg	tcaatgcggt	caatctggaa	gtgctgccgg	caccggaata	11760
cagctcggcg	cagcgcgcgc	tcagcgacgc	cgactggcat	ttcatcgccg	atcctgcgtc	11820
gaggttttac	aacctcgtct	tggctgattg	tggggccggc	ttcttcgacc	cgctgaccgc	11880
cggcgtgctg	tccacgggtg	ccgggtgtcgt	ggtcgtggca	agtgtctcaa	tcgacggcgc	11940
acaacaggcg	tgggtcgcgt	tggactgggt	gcgcaacaac	ggttaccaag	atgtggcgag	12000
ccgcgcacgc	gtggtcacga	atcacatcat	gccggggagaa	cccaatgtcg	cagttaaaga	12060
cctggtgcgg	catttcgaac	agcaagttca	accgggcggg	gtcgtgggtca	tgccgtggga	12120
caggcacatt	gcggccggaa	ccgagatttc	actcgacttg	ctcgacccta	tctacaagcg	12180
caaggctctc	gaattggccg	cagcgctatc	cgacgatttc	gagagggctg	gacgtcggtg	12240
agcgcacctg	ctggtgctgc	tggctctacc	gccgcggggg	caaccgctgc	gcggcctgcc	12300
accaccgggg	tgacgatcct	gaccggcaga	cggatgaccg	atgttggtact	gccagcggcg	12360
gtgccgatgg	aaacttatat	tgacgacacc	gtcgcgggtg	tttccgaggt	gttggaagac	12420
acgcgggctg	atgtactcgg	cggcttcgac	tttaccgcgc	aaggcgtgtg	ggcgttcgct	12480
cgctcccgat	cgccgcggct	gaagctcgac	cagtcactcg	atgacgcggg	ggtgggtcgac	12540
gggtcactgc	tgactctggt	gtcagtcagt	cgcaccgagc	gctaccgacc	gttggtcgag	12600
gatgtcatcg	acgcgactgc	cgtgcttgac	gagtcacctg	agttcgaccg	cacggcattg	12660
aatgcctttg	tgggggcggc	gatcccgcct	ttgaccgcgc	ccgtcatcgg	gatggcgatg	12720
cgggcggtgt	gggaaactgg	gcgtagcttg	tgggtggcgt	tggcgattgg	catcctgggg	12780
atcgctgtgc	tggtaggcag	cttcgtcgcg	aacaggttct	accagagcgg	ccacctggcc	12840
gagtgcctac	tggtcacgac	gtatctgctg	atcgcaaccg	ccgcagcgct	ggccgtgccc	12900

ttgccgcgcg	gggtcaactc	gttggggg	ccacaagttg	ccggcgccgc	tacggccgtg	12960
ctgtttttga	ccttgatgac	gcggggcg	cctcggaagc	gtcatgagtt	ggcgctcgtt	13020
gccgtgatca	ccgctatcgc	ggtcatcg	gccgcgctg	ccttcggcta	tggataccag	13080
gactgggtcc	ccgcggggg	gacgcattc	gggctgttca	ttgtgacgaa	tgcggccaag	13140
ctgaccgtcg	cggtcgcgcg	gatcgcgctg	ccgcccattc	cggtagccgg	cgaaaccgtg	13200
gacaacgagg	agttgctcga	tcccgctcgc	accccgagg	ctaccagcga	agaaaccccg	13260
acctggcagg	ccatcatcgc	gtcgggtgcc	gcgtccgcg	tccggctcac	cgagcgcagc	13320
aaactggcca	agcaacttct	gatcggatac	gtcacgtcgg	gcaccctgat	tctggctgcc	13380
ggtgccatcg	cggtcgtggg	gcgcggggc	ttctttgtac	acagcctggg	ggtcgcgggt	13440
ttgatcacga	ccgtctgcgg	atttcgctcg	cggctttacg	ccgagcgtg	gtgtgcgtgg	13500
gcgttgctgg	cggcgacggg	cgcgattccg	acgggtctga	cggccaaact	catcatctgg	13560
taccgcact	atgcctggct	gttgttgagc	gtctacctca	cggtagccct	ggttgcgctc	13620
gtggtggtcg	ggtcgatggc	tcacgtccgg	cgcgtttcac	cggtcgtaaa	acgaactctg	13680
gaattgatcg	acggcgccat	gatcgctgcc	atcattccca	tgtctgtgtg	gatcaccggg	13740
gtgtacgaca	cggtcgcaa	tatccggttc	tga			13773

&lt;210&gt; 3

&lt;211&gt; 3909

&lt;212&gt; DNA

&lt;213&gt; Mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; RD1-AP34 (a 3909 bp fragment of the M. tuberculosis H37Rv genome)

&lt;400&gt; 3

gaattcccat	ccagtgagtt	caagggtcaag	cggcgccccc	ctggccaggc	atttctcgtc	60
tcgccagacg	gcaaagaggt	catccaggcc	ccctacatcg	agcctccaga	agaagtgttc	120
gcagcacccc	caagcgccgg	ttaagattat	ttcattgccg	gtgtagcagg	acccgagctc	180
agcccggtaa	tgcagttcgg	gcaatgctga	ccatcggtt	tgtttcggc	tataaccgaa	240
cggtttggtg	acgggataca	aatacaggga	gggaagaagt	aggcaaattg	aaaaaatgtc	300
acatgatccg	atcgctgcgc	acattggcac	gcaagtgagc	gacaacgctc	tgcacggcgt	360
gacggccggc	tcgacggcgc	tgacgtcggg	gaccgggctg	gttcccgcgg	gggccgatga	420
ggtctccgcc	caagcggcga	cggcgttcac	atcggagggc	atccaattgc	tggttccaa	480
tgcatcggcc	caagaccagc	tccaccgtgc	gggcgaagcg	gtccaggacg	tcgcccgcac	540
ctattcgcaa	atcgacgacg	gcgcgcgcgg	cgtcttcgcc	gaataggccc	ccaacacatc	600
ggagggagtg	atcaccatgc	tgtggcacgc	aatgccaccg	gagctaaata	ccgcacggct	660
gatggccggc	gcgggtccgg	ctccaatgct	tgccggcgcc	gcgggatggc	agacgctttc	720
ggcggtctcg	gacgtccagg	ccgtcgagtt	gaccgcgcgc	ctgaactctc	tgggagaagc	780
ctggactgga	gggtggcagc	acaaggcgct	tgccggctga	acgccgatgg	tggtctggct	840
acaaaccgcg	tcaacacagg	ccaagaccgc	tgcatgacg	gcgacggcgc	aagccgcggc	900
atacacccag	gccatggcca	cgacgcgcgc	gctgccggag	atcgccgcca	accacatcac	960
ccaggccgct	cttacggcca	ccaacttctt	cggtatcaac	acgatcccga	tcgcgttgac	1020
cgagatggat	tatttcatcc	gtatgtggaa	ccaggcagcc	ctggcaatgg	aggtctacca	1080
ggccgagacc	gcggttaaca	cgcttttcga	gaagctcgag	ccgatggcgt	cgatccttga	1140
tcccggcgcg	agccagagca	cgacgaaccc	gatcttcgga	atgccctccc	ctggcagctc	1200
aacaccggtt	ggccagttgc	cgccggcgcc	taccagacc	ctcgccaac	tgggtgagat	1260
gagcgcccg	atgcagcagc	tgacccagcc	gctgcagcag	gtgacgtcgt	tgttcagcca	1320
ggtggggcgc	accggcgccg	gcaacccagc	cgacgaggaa	gccgcgcaga	tgggcctgct	1380
cggcaccagt	ccgctgtcga	accatccgct	ggctggtgga	tcaggcccca	gcgcgggcgc	1440
gggcctgctg	cgcgcggagt	cgctacctgg	cgcaggtggg	tcgttgaccc	gcacgcgcgt	1500
gatgtctcag	ctgatcgaaa	agccggttgc	cccctcgggt	atgccggcgg	ctgctgccgg	1560
atcgctcggc	acgggtggcg	ccgctccggt	gggtgcggga	gcgatggggc	aggggtgcga	1620
atccggcgcc	tccaccaggc	cgggtctggt	cgcgcggcca	ccgctcgcgc	aggagcgtga	1680
agaagacgac	gaggacgact	gggacgaaga	ggacgactgg	tgagctcccg	taatgacaac	1740
agacttccc	gccacccggg	ccggaagact	tgccaacatt	ttggcgagga	aggtaaagag	1800
agaaagttag	ccagcatggc	agagatgaag	accgatgccg	ctaccctcgc	gcaggaggca	1860
ggtaatttcg	agcggatctc	cggcgacctg	aaaaccgaga	tcgaccaggt	ggagtcgacg	1920
gcaggttcgt	tgcagggccca	gtggcgcgcc	gcggcgggga	cggccgccca	ggccgcgggtg	1980
gtgcgcttcc	aagaagcagc	caataagcag	aagcaggaac	tcgacgagat	ctcgacgaat	2040



attcgtcagg	cggcggtcca	atactcgagg	gccgacgagg	agcagcagca	ggcgctgtcc	2100
tcgcaaatgg	gcttctgacc	cgctaatacg	aaaagaaacg	gagcaaaaac	atgacagagc	2160
agcagtggaa	tttcgcgggt	atcgaggccg	cggcaagcgc	aatccaggga	aatgtcacgt	2220
ccattcattc	cctccttgac	gaggggaagc	agtccctgac	caagctcgca	gcggcctggg	2280
gcggtagcgg	ttcggaggcg	taccagggtg	tccagcaaaa	atgggacgcc	acggctaccg	2340
agctgaacaa	cgcgctgcag	aacctggcgc	ggacgatcag	cgaagccggg	caggcaatgg	2400
cttcgaccga	aggcaacgtc	actgggatgt	tcgcataggg	caacgccgag	ttcgcgtaga	2460
atagcgaaac	acgggatcgg	gcgagttcga	ccttcgcgtc	gtctcgcctt	ttctcgtgtt	2520
tatacgtttg	agcgcaactc	gagaggttgt	catggcgggc	gactacgaca	agctcttccg	2580
gccgcacgaa	ggtatggaag	ctccggacga	tatggcagcg	cagccgttct	tcgaccccg	2640
tgcttcgttt	cgcgcggcgc	ccgcacgcgc	aaacctaccg	aagcccaacg	gccagactcc	2700
gccccgcagc	tccgacgacc	tgtcggagcg	gttcgtgtcg	gccccgcgcg	cgccaccccc	2760
acccccacct	ccgcctccgc	caactccgat	gccgatcgcc	gcaggagagc	cgccctcgcc	2820
ggaaccggcc	gcacctaacc	cacccacacc	ccccatgccc	atcgccggac	ccgaaccggc	2880
ccccccaaa	ccacccacac	cccccatgcc	catcgccgga	cccgaaccgg	ccccacccaa	2940
accacccaca	cctccgatgc	ccatcgccgg	acctgcaccc	accccaaccg	aatcccagtt	3000
ggcgcccccc	agaccaccga	caccacaaac	gccaaaccga	gcgccgcagc	aaccggaatc	3060
accggcgccc	cacgtaccct	cgcacggggc	acatcaaccc	cggcgccacc	caccagcacc	3120
gccctgggca	aagatgccaa	tcggcgaaac	cccgcgccgt	ccgtccagac	cgtctgcgtc	3180
ccgggcccga	ccaccgaccc	ggcctgcccc	ccaacactcc	cgacgtgcgc	gccgggggtca	3240
ccgctatcgc	acagacaccg	aacgaaacgt	cgggaaggta	gcaactgggt	catccatcca	3300
ggcgcggtcg	cgggcagagg	aagcatccgg	cgcgcggtc	gcccccgga	cggagccctc	3360
gccagcgccg	ttgggccaac	cgagatcgta	tctggctccg	cccaccgcgc	ccgcgcgcgac	3420
agaacctccc	cccagcccc	cgccgcagcg	caactccggt	cggcgtgccc	agcgacgcgt	3480
ccaccccgat	ttagccgccc	aacatgccgc	ggcgcaacct	gattcaatta	cggccgcaac	3540
cactggcggt	cgtcgccgca	agcgtgcagc	gccggatctc	gacgcgacac	agaaatcctt	3600
aaggccggcg	gccaaggggc	cgaagggtgaa	gaagggtgaag	ccccagaaac	cgaaggccac	3660
gaagccgccc	aaagtgggtg	cgcagcgccg	ctggcgacat	tgggtgcatg	cgttgacgcg	3720
aatcaacctg	ggcctgtcac	ccgacgagaa	gtacgagctg	gacctgcacg	ctcgagtccg	3780
ccgcaatccc	cgcgggtcgt	atcagatcgc	cgtcgtcggt	ctcaaagggt	gggctggcaa	3840
aaccacgctg	acagcagcgt	tggggctcgac	gttggctcag	gtgcggggcc	accggatcct	3900
ggctctaga						3909

&lt;210&gt; 4

&lt;211&gt; 324

&lt;212&gt; DNA

&lt;213&gt; Mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; DNA sequence Rv3861

&lt;400&gt; 4

gtgacctggc	tggctgaccc	ggtcggcaac	agcaggatcg	cccagcgca	ggcctgcaaa	60
acgtcaatct	cggcgcccat	cgtcgaatcc	tggcgggcgc	aacgcggcgc	gcaatgtgga	120
cagcgcgaga	aatcttgcgt	atgttctcgc	gctgtccaca	tccagggcat	ctcaccgcca	180
ctgttccgca	gacccctcga	accagcggtc	caggcgccgg	ttgcgtcatg	ccgattgggc	240
agacaccg	tggtcgcgca	ccgggtaacc	gttgcgctcg	gccagggatc	gcagctggcc	300
caacgcgaat	gccccgcgcc	ggcc				324

&lt;210&gt; 5

&lt;211&gt; 348

&lt;212&gt; DNA

&lt;213&gt; Mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; DNA sequence Rv3862c-whiB6

&lt;400&gt; 5

atgcgatacg	ctttcgcggc	agaggctaca	acctgcaacg	ctttctggag	gaacgtagac	60
------------	------------	------------	------------	------------	------------	----

```

atgacagtaa cgcacctgta tgagggtccc ctcggcggtt gcacgcaaga tcccgatcgt 120
tgagcagcga ctcccgacga cgaggccaag accctgtgcc gggcttgccc gcgccggtgg 180
ctgtgtgcac gcgacgccgt cgagtccgcg ggtgcggaag ggctgtgggc aggggtcgta 240
attcccgaat caggccgggc gcgggcattc gcgttgggccc agctgcgata cctggccgag 300
cgcaacgggt acccgggtgcg cgaccaccgc gtgtctgccc aatcggca 348

```

<210> 6

<211> 1176

<212> DNA

<213> *Mycobacterium tuberculosis*

<220>

<223> DNA sequence RV3863

<400> 6

```

atggcggggc agcggaaagt ctgcccaccg tcccggctag taccgcgcaa taagggatca 60
acgcagatgt ctaaagcagg gtcgactgtc ggaccggcgc cgctggtcgc gtgcagcggc 120
ggcacatcag acgtgattga gccccgtcgc ggtgtcgcga tcattggcca ctctgtccga 180
gtcggcacc cagatcgacga ttctcgaatc tctcagacac atctgcgagc ggtatccgat 240
gatggacggg ggcggatcgt cggcaacatc ccgagaggta tgttcgtcgc cggacgacgc 300
ggcagctcgg tgaccgtcag cgataagacc ctaatccgat tcggcgatcc ccctggaggc 360
aaggcggtga cgttcgaagt cgtcaggccg tcggattccg ctgcacagca cggccgcgta 420
caaccatcag cggacctgtc ggacgaccgc gcgcacaacg ctgcgccggt cgcaccggac 480
cccggcgtgg ttgcgcgagg ggcggccgcg gctgcgcgcc gtcgtgaact tgacatcagc 540
caacgcagct tggcggccga cgggatcatc aacgcgggcg cgtcatcgc gttcgagaaa 600
ggccgtagtt ggcggccgga acggaccgcg gcaaaactcg aagaagtgtc gcagtggccc 660
gctggaacca tcgcgcgaat ccgtcggggc gagcccaccg agcccgcaac aaaccccgac 720
gcgtcccccg gactccggcc tgccgacggc ccggcgctct tgatcgcgca ggctgtcacc 780
gcccgcgtag acggtgtcag tctggctatc gcagcgttgc cggcgaccga ggaccccgag 840
ttcaccgaac gtgcgcgcgc gatccttgct gatttgcgc agctcgaggc gattgcgctc 900
caagcaacc gcacagccg gattaccgc gaattgatca aggcgttggg cgcgggtacgt 960
cgccaccagc acgaattaat gaggttggga gcaaccgccc ctggtgccac actggcgagc 1020
cgcttatatg ccgcacggcg gcgcgcgaac ctttccaccc tggagactgc ccaagcggcc 1080
ggcgtcgcag aagaaatgat cgtcggcgcc gaagccgagg aagagttgcc agccgaggcc 1140
accgaagcga tcgaagcact gatccgtcag atcaat 1176

```

<210> 7

<211> 1206

<212> DNA

<213> *Mycobacterium tuberculosis*

<220>

<223> DNA sequence Rv3864

<400> 7

```

atggcatcgg gtagcgggtc ttgcaagacg acgagtaact ttatttgggg ccagttactc 60
ttgcttggag agggaatccc cgacccaggc gacattttca acaccggttc gtcgctgttc 120
aaacaaatca gcgacaaaat gggactcgcc attccgggca ccaactggat cggccaagcg 180
gcggaagctt acctaaacca gaacatcgcg caacaacttc gcgcacaggt gatgggcgat 240
ctcgacaaat taaccggcaa catgatctcg aatcaggcca aatacgtctc cgatacgcgc 300
gacgtcctgc gggccatgaa gaagatgatt gacgggtgtc acaaggtttg taagggcctc 360
gaaaagattc cgtgtctcgc ccacttgtgg tcgtgggagc tcgcaatccc tatgtccggc 420
atcgcgatgg ccgttgtcgc cggcgcattg ctctatctaa cgattatgac gctgatgaat 480
gcgaccaacc tgaggggaat tctcggcagg ctgatcgaga tgttgacgac cttgccaaag 540
ttccccggcc tgccccgggt gccccagcctg cccgacatca tcgacggcct ctggccgccc 600
aagttgcccc acattccgat ccccggcctg cccgacatcc cgggcctacc cgacttcaaa 660
tggccgcccc cccccggcag cccgttggtc cccgacctcc cgtcgttccc agggttcccc 720
gggttccccg agttccccgc catccccggg ttccccgcac tgccccgggt gccccagcatt 780
cccaacttgt tccccggctt gccgggtctg ggcgacctgc tgccccggct aggcgatttg 840

```

```

ggcaagttac ccacctggac tgagctggcc gctttgcctg acttcttggg cggcttcgcc 900
ggcctgcccc gcttggggtt tggcaatctg ctcagctttg ccagtttgcc caccgtgggt 960
caggtgaccg ccacctggg tcagctgcaa cagctcgtgg cggccggcgg tggccccagc 1020
caactggcca gcatgggcag ccaacaagcg caactgatct cgtcgcaggc ccagcaagga 1080
ggccagcagc acgccaccct cgtgagcgac aagaaggaag acgaggaagg cgtggccgag 1140
gcgagcgtg caccatcga cgctggcacc gcggccagcc aacgggggca ggaggggacc 1200
gtcctt 1206

```

<210> 8

<211> 309

<212> DNA

<213> Mycobacterium tuberculosis

<220>

<223> DNA sequence Rv3865

<400> 8

```

atgaccggat ttctcggtgt cgtgccttcg ttcctgaagg tgctggcggg catgcacaac 60
gagatcgtgg gtgatataaa aagggcgacc gatacggtcg ccgggattag cggacgagtt 120
cagcttaccg atggttcggt cagctcgaaa ttcaatgaca cgctgcaaga gtttgagacc 180
acccgtagca gcacgggcac gggtttgtag ggagtcacca gcggactggc caataatctg 240
ctcgcagccg ccggcgccca cctcaaggcc gacgatggcc tagccggtgt tatcgacaag 300
attttcgggt 309

```

<210> 9

<211> 849

<212> DNA

<213> Mycobacterium tuberculosis

<220>

<223> DNA sequence Rv3866

<400> 9

```

atgacgggtc cgtccgctgc aggccgcgcg ggcaccgccg acaacgtggt cggcgtcgag 60
gtaaccatcg acggcatggt ggtgatcgcc gatcggttac acctggttga ttccctgtc 120
acgcttggga ttccggccga tatcccgcaa gaggatctgc gagacatcgt ctgggaacag 180
gtgcagcgtg acctcacagc gcaaggggtg ctgcacctcc acggggagcc ccaaccgacg 240
gtcgcggaga tggtcgaaac cctgggcagg ccagatcgga ccttggaggg tcgctggtgg 300
cggcgcgaca ttggcgccgt catggtgcgc ttcgtcgtgt gccgcagggg cgaccgccat 360
gtgatcgcg cgcgcgacgg cgacatgctg gtgctgcagt tgggtggcgc gcaggtcggc 420
ttggcgggca tgggtgacagc ggtgctgggg cccgccgaac ccgccaacgt cgaacccctg 480
acgggtgtgg caaccgagct agccgaatgc acaaccgcgt cccaattgac gcaatacggg 540
atcgcaccgg cctcggcccc cgtctatgcc gagatcgtgg gtaaccgcac cggctggggtg 600
gagatcgttg ccagccaacg ccaccccggc ggcaccacga cgcagaccga cgccgcgcgt 660
ggcgtccttg actccaagct cggtaggctg gtgtcgcttc cccgccgtgt tggaggcgac 720
ctgtacggaa gcttctctgc cggcactcag cagaacttgg agcgtgcgct ggacggcttg 780
ctagagctgc tccctgcggg cgcttggcta gatcacacct cagatcacgc acaagcctcc 840
tcccagggc 849

```

<210> 10

<211> 552

<212> DNA

<213> mycobacterium tuberculosis

<220>

<223> DNA sequence Rv3867

<400> 10

atggtggacc	cgccgggcaa	cgacgacgac	cacggtgatc	tcgacgccct	cgattttctcc	60
gccgcccaca	ccaacgaggc	gtcgccgctg	gacgccttag	acgactatgc	gccggtgcag	120
accgatgacg	ccgaaggcga	cctggacgcc	ctccatgcgc	tcaccgaacg	cgacgaggag	180
ccggagctgg	agttgttcac	ggtagccaac	cctcaagggt	cggtgtcggg	ctcaaccctg	240
atggacggca	gaatccagca	cgtcgagctg	acggacaagg	cgaccagcat	gtccgaagcg	300
cagctggccg	acgagatctt	cgttattgcc	gatctggccc	gccaaaaggc	gcgggcgtcg	360
cagtacacgt	tcattggtgga	gaacatcggt	gaactgaccg	acgaagacgc	agaaggcagc	420
gccctgctgc	gggaattcgt	ggggatgacc	ctgaatctgc	cgacgccgga	agaggctgcc	480
gcagccgaag	ccgaagtgtt	cgccacccgc	tacgatgtcg	actacacctc	ccggtacaag	540
gccgatgact	ga					552

&lt;210&gt; 11

&lt;211&gt; 1722

&lt;212&gt; DNA

&lt;213&gt; mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; DNA sequence Rv3868

&lt;400&gt; 11

atgactgata	gcttggccag	tctgttcgaa	agcgccgtca	gcatgttgcc	gatgtcggag	60
gcgcggctcg	tagatctgtt	caccgagatc	accaaactacg	acgaatccgc	ttgcgacgca	120
tggatcggcc	ggatccgggtg	tggggacacc	gaccgggtga	cgctgtttcg	cgccctggtat	180
tcgcgccgca	atttcggaca	gttgtcggga	tcgggtccaga	tctcgatgag	cacgttaaac	240
gccaggattg	ccatcggggg	gctgtacggc	gatatcacct	acccggtcac	ctcgccgcta	300
gcgatcacca	tgggctttgc	cgcatgcbag	gcagcgcaag	gcaattacgc	cgacgccatg	360
gaggccttag	aggccgcccc	ggtcgcgggt	tcgagacacc	tggtggcgtg	gatgaaggcg	420
gttgtctacg	gcgcggccga	acgctggacc	gacgtgatcg	accagggtcaa	gagtgtctgg	480
aaatggcccg	acaagttttt	ggccggcgcg	gccggtgtgg	cgacgggggt	tgccgcggca	540
aacctggcct	tggttcaccga	agccgaacgc	cgactcaccg	aggccaacga	ctcgcccgcc	600
ggtagggcgt	gtgcgcgcgc	catcgccctg	tatctggcga	tggtcacggc	cagccagggc	660
aacgaaagcg	cgcggtgggc	gctgctggaa	tggttacaga	ccactcacc	cgagcccaaa	720
gtggctgcgg	cgctgaagga	tcctctctac	cggtgaaga	cgaccaccgc	cgaacagatc	780
gcatcccgcg	ccgatccctg	ggatccgggc	agtgtcgtga	ccgacaactc	cgcccgggag	840
cggtctctcg	ccgagggcca	agccgaactc	gaccgccaaa	ttgggtctac	ccgggttaaa	900
aatcagattg	aacgctaccg	cgcggcgacg	ctgatggccc	gggtccgcgc	cgccaagggt	960
atgaaggctg	cccagcccag	caagcacatg	atcttcaccg	gaccgcccgg	taccggcaag	1020
accacgatcg	cgcggttggt	ggccaatatc	ctggccggct	taggcgtcat	tgccgaacct	1080
aaactcgtcg	agacgtcgcg	caaggacttc	gtcgccgagt	acgaggggca	atcggcgggtc	1140
aagaccgcta	agacgatcga	tcaggcgctg	ggcggggtgc	ttttcatcga	cgaggcttat	1200
gcgctggtgc	aggaaagaga	cggccgcacc	gatccgttcg	gtcaagaggc	gctggacacg	1260
ctgctggcgc	ggatggagaa	cgaccgggac	cggtggtgg	tgatcatcgc	cggttacagc	1320
tccgacatag	atcggtctgt	ggaaaccaac	gagggctctg	ggtcgcgggt	cgccactcgc	1380
atcgagttcg	acacctattc	ccccgaggaa	ctcctcgaga	tcgccaacgt	cattgccgct	1440
gctgatgatt	cggcgttgac	cgagaggcg	gccgagaact	ttcttcaggc	cgccaagcag	1500
ttggagcagc	gcatgttcg	cggccggcgc	gccctggacg	tcgccggcaa	cggtcggtat	1560
gcgcgccagc	tggtggaggc	cagcgagcaa	tgccgggaca	tgctctagc	ccaggtcctc	1620
gatatcgaca	cctcgacga	agaccggctt	cgcgagatca	acggctcaga	tatggcggag	1680
gctatcgccg	cgggtgcacgc	acacctcaac	atgagagaat	ga		1722

&lt;210&gt; 12

&lt;211&gt; 1443

&lt;212&gt; DNA

&lt;213&gt; mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; DNA sequence Rv3869

&lt;400&gt; 12

atggggcttc	gcctcaccac	caaggttcag	gttagcggct	ggcgttttct	gctgcgccgg	60
ctcgaacacg	ccatcgtgcg	ccgggacacc	cggatgtttg	acgacccgct	gcagttctac	120
agccgctcga	tcgctcttgg	catcgtcgtc	gcggtcctga	ttctggcggg	tgccgcgctg	180
ctggcgtaact	tcaaaccaca	aggcaaactc	ggcggcacca	gcctgttcac	cgaccgcgcg	240
accaaccagc	tttacgtgct	gctgtccgga	cagttgcatc	cggtctacaa	cctgacttcg	300
gcgcggctgg	tgctgggcaa	tccggccaac	ccggccaccg	tgaagtccct	cgaactgagc	360
aagctgccga	tgggcccagac	cggttgaatc	cccggcgccc	cctacgccac	gcctgtttcg	420
gcgggcagca	cctcgatctg	gaccctatgc	gacaccgtcg	cccagagccga	ctccacttcc	480
ccggtagtgc	agaccgcggt	catcgcgatg	ccgttggaga	tcgatgcttc	gatcgatccg	540
ctccagtcac	acgaagcggg	gctggtgtcc	taccagggcg	aaacctggat	cgtcacaact	600
aagggacgcc	acgccataga	tctgaccgac	cgcgccctca	cctcgtcgat	ggggataccg	660
gtgacggcca	ggccaacccc	gatctcggag	ggcatgttca	acgcgctgcc	tgatatgggg	720
ccctggcagc	tgccgcgcat	accggcgggc	ggcgcgccca	attcgcttgg	cctacctgat	780
gatctagtga	tcggatcggg	cttccagatc	cacaccgaca	agggcccgca	atactatgtg	840
gtgctgcccg	acggcatcgc	gcaggtcaac	gcgacaaccg	ctgcggcgct	gcgcgccacc	900
caggcgcacg	ggctggtcgc	gccaccggca	atggtgccca	gtctggctcg	cagaatcgcc	960
gaacgggtat	acccctcacc	gctacccgat	gaaccgctca	agatcgtgtc	ccggccgcag	1020
gatcccgcg	tgtgctggtc	atggcaacgc	agcgcggcg	accagtcgcc	gcagtcaacg	1080
gtgctgtccg	gccggcatct	gccgatatcg	ccctcagcga	tgaacatggg	gatcaagcag	1140
atccacggga	cggcgaccgt	ttacctcgac	ggcggaata	tcgtggcact	gcaatcccc	1200
gacccctgat	acaccgaatc	gatgtactac	atcgatccac	agggcgctgc	ttatgggggtg	1260
cctaaccgcg	agacagccaa	gtcgctgggc	ctgagttcac	cccaaaacgc	gccctgggag	1320
atcgttcgct	tcctggtcga	cggtcgggtg	ctgtcgaaag	atgccgcact	gctcgagcac	1380
gacacgctgc	ccgctgaccc	tagccccga	aaagttcccg	ccggagccct	cgagagcccc	1440
tga						1443

&lt;210&gt; 13

&lt;211&gt; 2244

&lt;212&gt; DNA

&lt;213&gt; mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; DNA sequence Rv3870

&lt;400&gt; 13

atgacgacca	agaagtccac	tcccaccatt	accggtggcc	cccgggtgac	cccgggagag	60
atcagcctca	cgccgcccga	tgacctgggc	atcgacatcc	caccgtcggg	cgccccaaag	120
atccttccct	acgtgatggg	tggcgccatg	ctcggcatag	tcgccatcat	ggtggccggc	180
ggcaccaggc	agctgtcgcc	gtacatgttg	atgatgccgc	tgatgatgat	cgtgatgatg	240
gtcggcggtc	tggccggtag	caccggtggt	ggcggcaaga	aggtgcccga	aatcaacgcc	300
gaccgcaagg	agtacctgcg	gtatttggca	ggactacgca	cccagtgac	gtcctcggcc	360
acctctcagg	tggcgttctt	ctcctaccac	gcaccgcate	ccgaggatct	gttgtcgatc	420
gtcggcaccc	aacggcagtg	gtcccggccg	gccaacgcgc	acttctatgc	ggccacccga	480
atcggtatcg	gtgaccagcc	ggcgggtggat	cgattattga	agccggccgt	cggcggggag	540
ttggccgccc	ccagcgcagc	acctcagccg	ttcctggagc	cggtcagtca	tatgtgggtg	600
gtcaagtttc	tacgaaccca	tggattgatc	catgactgcc	cgaaactgct	gcaactccgt	660
acctttccga	ctatcgcgat	cggcggggac	ttggcggggg	cagccggcct	gatgacggcg	720
atgatctgtc	acctagccgt	gttccacca	ccggacctgc	tgcatatccg	ggtgctcacc	780
gaggaacccg	acgaccccga	ctggtcctgg	ctcaaattgg	ttccgcacgt	acagcaccag	840
accgaaaccg	atgcggccgg	gtccaccggg	ctgatcttca	cgcgccagga	aggtctgtcg	900
gacctggccg	cgcgcggggc	acacgcaccc	gattecgttc	ccggcgggcc	ctacgtagtc	960
gtcgtcgacc	tgaccggcgg	caaggctgga	ttcccggccc	acggtagggc	cggtgtcacg	1020
gtgatcacgt	tgggcaacca	tcgcggctcg	gcctaccgca	tcagggtgca	cgaggatggg	1080
acggctgatg	accggctccc	taaccaatcg	tttcgccagg	tgacatcggt	caccgatccg	1140
atgtcgccgc	agcaagccag	ccgtatcgcg	cgaaagttgg	ccggatggtc	catcacgggc	1200
accatcctcg	acaagagctc	gcgggtccag	aagaaggtgg	ccaccgactg	gcaccagctg	1260
gtcgggtgcg	aaagtgtcga	ggagataaca	cttcccgcgt	ggaggatgta	caccgacacc	1320
gaccgtgacc	ggctaagaat	cccgtttggt	catgaactaa	agaccggcaa	cgatcatgtac	1380
ctggacatca	aagagggcgc	ggaattcggc	gccggaccgc	acggcatgct	catcggggacc	1440
acgggggtctg	ggaagtccga	attcctgcgc	acctgatcc	tgctcgtggt	ggcaatgact	1500

catccagatc	aggtgaatct	cctgctcacc	gacttcaaag	gtggttcaac	cttcctggga	1560
atggaaaaagc	ttccgcacac	tgccgctgtc	gtcaccaaca	tggccgagga	agccgagctc	1620
gtcagccgga	tgggcgaggt	gttgaccgga	gaactcgatc	ggcgcagtc	gacccctccga	1680
caggccggga	tgaaagtccg	cgcggccgga	gccctgtccg	gcgtggccga	atacgagaag	1740
taccgcgaac	gcggtgccga	cctacccccc	ctgccaacgc	ttttcgtcgt	cgtcgacgag	1800
ttcgcgagc	tgttgcagag	tcacccggac	ttcatcgggc	tgttcgaccg	gatctgccgc	1860
gtcggggcgt	cgctgagggt	ccatctgctg	ctggctaccc	agtcgctgca	gaccggcgggt	1920
gttcgcatcg	acaaactgga	gccaaacctg	acatatcgaa	tcgcattgctg	caccaccagc	1980
tctcatgaat	ccaaggcgggt	aatcggcaca	ccggaggcgc	agtacatcac	caacaaggag	2040
agcgggtgtcg	ggttttctccg	ggtcggcatg	gaagaccccg	tcaagttcag	caccttctac	2100
atcagtgggc	catacatgcc	gccggcggca	ggcgtcgaaa	ccaatggtga	agccggaggg	2160
cccgggtcaac	agaccactag	acaagccgcg	cgcattcaca	ggttcaccgc	ggcaccgggt	2220
ctcgaggagg	cgccgacacc	gtga				2244

&lt;210&gt; 14

&lt;211&gt; 1776

&lt;212&gt; DNA

&lt;213&gt; mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; DNA sequence Rv3871

&lt;400&gt; 14

atgactgctg	aaccggaagt	acggacgctg	cgcgagggtg	tgctggacca	gctcggcact	60
gctgaatcgc	gtgcgtacaa	gatgtggctg	ccgccgttga	ccaatccgggt	cccgtcacaac	120
gagctcatcg	cccgtgatcg	gcgacaaccc	ctgcgatttg	ccctggggat	catggatgaa	180
ccgcgccgcc	atctacagga	tgtgtggggc	gtagacgttt	ccggggccgg	cggcaacatc	240
ggatttgggg	gcgcacctca	aaccgggaag	tcgacgctac	tgacagacgat	ggtgatgtcg	300
gccgccgcca	cacactcacc	gcgcaacggt	cagttctatt	gcacgacact	aggtggcggc	360
gggctgatct	atctcgaaaa	ccttccacac	gtcgggtggg	tagccaatcg	gtccgagccc	420
gacaaggta	accgggtgggt	cgcagagatg	caagccgtca	tgccgcaacg	ggaaaccacc	480
ttcaaggaa	accgagtggtg	ctcgatcggg	atgtaccggc	agctgcgtga	cgatccaagt	540
caaccggtg	cgtccgatcc	atacggcgac	gtctttctga	tcacgacgag	atggcccgggt	600
tttgcggcg	agttccccga	ccttgagggg	cagggttcaag	atctggccgc	ccaggggctg	660
gcgttcggcg	tccacgtcat	catctccacg	ccacgctgga	cagagctgaa	gtcgcgtggt	720
cgcgactacc	tcggcaccaa	gatcgagttc	cggcttggtg	acgtcaatga	aaccagatc	780
gaccggatta	cccgcgagat	cccggcgaat	cgtccgggtc	gggcagtgtc	gatggaaaag	840
caccatctga	tgatcggcgt	gcccagggtc	gacggcgtgc	acagcgccga	taacctggtg	900
gaggcgatca	ccgcgggggt	gacgcagatc	gcttcccagc	acaccgaaca	ggcacctccg	960
gtgcgggtcc	tgccggagcg	tatccacctg	cacgaactcg	acccgaaccc	gccgggacca	1020
gagtcgact	accgcactcg	ctgggagatt	ccgatcgggt	tgccgcgagac	ggacctgacg	1080
ccggctcact	gccacatgca	cacgaaccgc	cacctactga	tcttcgggtg	ggccaaatcg	1140
ggcaagacga	ccattgcccc	cgcgatcgcg	cgcgccattt	gtgcccga	cagtccccag	1200
cagggtgcgt	tcattgctgc	ggactaccgc	tcgggcctgc	tggacgcgggt	gccggacacc	1260
catctgctgg	gcgcgggcgc	gatcaaccgc	aacagcgcgt	cgctagacga	ggccgttcaa	1320
gcactggcgg	tcaacctgaa	gaagcgggtg	ccgccgaccg	acctgacgac	ggcgcagcta	1380
cgctcgcgtt	cgtgggtggag	cggatttgac	gtcgtgcttc	tggtcgacga	ttggcacatg	1440
atcgtgggtg	ccgcgggggg	gatgccgccc	atggcacccg	tgcccccggt	attgccggcg	1500
gcggcagata	tcgggttgca	catcattgtc	acctgtcaga	tgagccaggc	ttacaaggca	1560
accatggaca	agttcgtcgg	cgcgcatttc	gggtcggggc	ctccgacaat	gttcctttcg	1620
ggcgagaagc	aggaattccc	atccagtgag	ttcaagggtca	agcggcgccc	ccctggccag	1680
gcattttctg	tctcgccaga	cggcaaagag	gtcatccagg	ccccctacat	cgagcctcca	1740
gaagaagtgt	tcgcagcacc	cccaagcgcc	ggttaa			1776

&lt;210&gt; 15

&lt;211&gt; 297

&lt;212&gt; DNA

&lt;213&gt; Mycobacterium tuberculosis



&lt;220&gt;

&lt;223&gt; PE coding sequence (Rv3872)

&lt;400&gt; 15

atggaaaaaa	tgtcacatga	tccgatcgct	gccgacattg	gcacgcaagt	gagcgacaac	60
gctctgcacg	gcgtgacggc	cggctcgacg	gcgctgacgt	cggtgaccgg	gctggttccc	120
gcggggggccg	atgaggtctc	cgcccaagcg	gcgacggcgt	tcacatcgga	gggcatccaa	180
ttgctggctt	ccaatgcata	ggcccaagac	cagctccacc	gtgcgggcga	agcgggtccag	240
gacgtcgcgc	gcacctattc	gcaaatcgac	gacggcgccg	ccggcgctct	cgccgaa	297

&lt;210&gt; 16

&lt;211&gt; 1104

&lt;212&gt; DNA

&lt;213&gt; Mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; PPE coding sequence (Rv3873)

&lt;400&gt; 16

atgctgtggc	acgcaatgcc	accggagcta	aataccgcac	ggctgatggc	cggcgcgggg	60
ccggctccaa	tgcttgccgc	ggccgcggga	tggcagacgc	tttcggcggc	tctggacgct	120
caggccgtcg	agttgaccgc	gcgcctgaac	tctctgggag	aagcctggac	tggaggtggc	180
agcgacaagg	cgttgccggc	tgcaacgccg	atggtggtct	ggctacaaac	cgctcaaca	240
caggccaaga	cccgtgcgat	gcaggcgacg	gcgcaagccg	cggcatacac	ccaggccatg	300
gccacgacgc	cgtegetgcc	ggagatcgcc	gccaaccaca	tcaccaggc	cgctccttacg	360
gccaccaact	tcttcggtat	caacacgatc	ccgatcgctg	tgaccgagat	ggattatttc	420
atccgtatgt	ggaaccaggc	agccctggca	atggaggtct	accaggccga	gaccgcggtt	480
aacacgcttt	tcgagaagct	cgagccgatg	gcgtcgatcc	ttgatcccgg	cgcgagccag	540
agcacgacga	accgatctt	cggaatgccc	tcccctggca	gctcaacacc	ggttggccag	600
ttgccgcggg	cggctaccca	gaccctcggc	caactgggtg	agatgagcgg	cccgatgcag	660
cagctgaccc	agccgctgca	gcaggtgacg	tcgttgttca	gccaggtggg	cggcaccggc	720
ggcgggaacc	cagccgacga	ggaagccgcg	cagatgggcc	tgctcggcac	cagtccgctg	780
tcgaaccatc	cgctgggtgg	tggatcaggc	cccagcgccg	gcgcgggcct	gctgcgcgcg	840
gagtcgctac	ctggcgccagg	tgggtcggtg	acccgcacgc	cgctgatgtc	tcagctgata	900
gaaaagccgg	ttgccccctc	ggtgatgccg	gcggctgctg	ccggatcgtc	ggcgacgggt	960
ggcgccgctc	cggtgggtgc	gggagcgatg	ggccaggggtg	cgcaatccgg	cggctccacc	1020
aggccgggtc	tggtcgcgcc	ggcaccgctc	gcgcaggagc	gtgaagaaga	cgacgaggac	1080
gactgggacg	aagaggacga	ctgg				1104

&lt;210&gt; 17

&lt;211&gt; 300

&lt;212&gt; DNA

&lt;213&gt; Mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; CFP-10 coding sequence (Rv3874)

&lt;400&gt; 17

atggcagaga	tgaagaccga	tgccgctacc	ctcgcgcagg	aggcaggtaa	tttcgagcgg	60
atctccggcg	acctgaaaac	ccagatcgac	caggtggagt	cgacggcagg	ttcgttgacg	120
ggccagtggc	gcggcgccgc	ggggacggcc	gcccaggccg	cgggtggtgcg	cttccaagaa	180
gcagccaata	agcagaagca	ggaactcgac	gagatctcga	cgaatattcg	tcaggccggc	240
gtccaatact	cgagggccga	cgaggagcag	cagcaggcgc	tgtcctcgca	aatgggcttc	300

&lt;210&gt; 18

&lt;211&gt; 285

&lt;212&gt; DNA

&lt;213&gt; Mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; ESAT-6 coding sequence (Rv3875)

&lt;400&gt; 18

atgacagagc	agcagtggaa	tttcgcggggt	atcgaggccg	cggcaagcgc	aatccaggga	60
aatgtcacgt	ccattcattc	cctccttgac	gaggggaagc	agtccctgac	caagctcgca	120
gcggcctggg	gcggtagcgg	ttcggaggcg	taccagggtg	tccagcaaaa	atgggacgcc	180
acggctaccg	agctgaacaa	cgcgtgcag	aacctggcgc	ggacgatcag	cgaagccggg	240
caggcaatgg	cttcgaccga	aggcaacgtc	actgggatgt	tcgca		285

&lt;210&gt; 19

&lt;211&gt; 2001

&lt;212&gt; DNA

&lt;213&gt; mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; DNA sequence Rv3876

&lt;400&gt; 19

atggcgggcg	actacgacaa	gctcttccgg	ccgcacgaag	gatatggaagc	tccggacgat	60
atggcagcgc	agccgttctt	cgaccccagt	gcttcgtttc	cgccggcgcc	cgcatcggca	120
aacctaccga	agcccaacgg	ccagactccg	cccccgacgt	ccgacgacct	gtcggagcgg	180
ttcgtgtcgg	ccccgcggcc	gccaccccc	ccccacctc	cgctccggcc	aactccgatg	240
ccgatcgccg	caggagagcc	gcctcgccg	gaaccggcgg	catctaaacc	acccacaccc	300
cccatgcccc	tcgcccggacc	cgaaccggcc	ccacccaaac	cacccacacc	ccccatgccc	360
atcgccggac	ccgaaccggc	cccacccaaa	ccacccacac	ctccgatgcc	catcgccgga	420
cctgcaccca	ccccaaccca	atcccagttg	gcgcccccca	gaccaccgac	accacaaacg	480
ccaaccggag	cgccgcagca	accggaatca	ccggcgcccc	acgtaccctc	gcacggggcca	540
catcaacccc	ggcgcaccgc	accagcaccg	ccctggggcaa	agatgccaat	cggcgaaccc	600
ccgcccgttc	cgtccagacc	gtctgcgtcc	ccggccgaac	caccgacccg	gcctgcccc	660
caacactccc	gacgtgcgcg	ccgggggtcac	cgctatcgca	cagacaccga	acgaaacgtc	720
gggaaggtag	caactgggtc	atccatccag	gcgcgggtgc	gggcagagga	agcatccggc	780
gcgcagctcg	cccccggaac	ggagccctcg	ccagcgccgt	tgggccaacc	gagatcgtat	840
ctgggtccgc	ccacccgccc	cgcgcgcaca	gaacctcccc	ccagccccctc	gccgcagcgc	900
aactccggtc	ggcgtgccga	gcgacgcgtc	caccccgatt	tagccgccca	acatgccgcg	960
gcgcaacctg	attcaattac	ggccgcaacc	actggcggtc	gtcgcgcgca	gcgtgcagcg	1020
ccgatctctg	acgcgacaca	gaaatcctta	aggccggcgg	ccaagggggc	gaaggtgaag	1080
aaggtgaagc	cccagaaacc	gaaggccacg	aagccgcccc	aagtgggtgc	gcagcgcggc	1140
tggcgacatt	gggtgcatgc	gttgacgcga	atcaacctgg	gcctgtcacc	cgacgagaag	1200
tacgagctgg	acctgcacgc	tcgagtcocg	cgcaatcccc	gcgggtcgta	tcagatcgcc	1260
gtcgtcggtc	tcaaagggtg	ggctggcaaa	accacgctga	cagcagcggt	ggggtcgacg	1320
ttggctcagg	tgcggggccga	ccggatcctg	gctctagacg	cggatccagg	cgccggaaac	1380
ctcgccgatac	gggtaggggcg	acaatcgggc	gcgaccatcg	ctgatgtgct	tgcagaaaaa	1440
gagctgtcgc	actacaacga	catccgcgca	cacactagcg	tcaatgcggt	caatctggaa	1500
gtgctgccgg	caccggaata	cagctcggcg	cagcgcgcgc	tcagcgacgc	cgactggcat	1560
ttcatcgccg	atcctgcgtc	gaggttttac	aacctcgtct	tggctgattg	tggggccggc	1620
ttcttcgacc	cgtgaccccg	cggcgtgctg	tccacgggtg	ccggtgtcgt	ggcgtgggca	1680
agtgtctcaa	tcgacggcgc	acaacaggcg	tcggtcgcgt	tggactgggt	gcgcaacaac	1740
ggttaccaag	atttggcgag	ccgcgcacgc	gtggtcatca	atcacatcat	gccggggagaa	1800
cccaatgtcg	cagttaaaga	cctggtcgcg	catttcgaac	agcaagttca	acccggccgg	1860
gtcgtgggtc	tgcggtggga	caggcacatt	gcggccggaa	ccgagatttc	actcgacttg	1920
ctcgacccta	tctacaagcg	caaggtcctc	gaattggccg	cagcgctatc	cgacgatttc	1980
gagagggctg	gacgtcgttg	a				2001

&lt;210&gt; 20

&lt;211&gt; 1536

&lt;212&gt; DNA

&lt;213&gt; mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; DNA sequence Rv3877

&lt;400&gt; 20

ttgagcgcac	ctgctggttc	tgctggtcct	accgcgcg	gggcaaccgc	tgcgcgccct	60
gccaccaccc	gggtgacgat	cctgaccggc	agacggatga	ccgatttggg	actgccagcg	120
gcggtgccga	tggaaactta	tattgacgac	accgtcgcg	tgctttccga	ggtgttgga	180
gacacgccgg	ctgatgtact	cgcgggcttc	gactttaccg	cgcaaggcgt	gtgggcgttc	240
gctcgtcccg	gacgcccgc	gctgaagctc	gaccagtcac	tcgatgacgc	cggggtgggc	300
gacgggtcac	tgctgactct	ggtgtcagtc	agtcgcaccg	agcgtaccg	accgttgggc	360
gaggatgtca	tcgacgcgat	cgccgtgctt	gacgagtcac	ctgagttcga	ccgcacggca	420
ttgaatcgct	ttgtgggggc	ggcgatcccg	cttttgaccg	cgcccgatcat	cgggatggcg	480
atgcgggcgt	ggtgggaaac	tgggcgtagc	ttgtgggtgg	cgttggcgat	tggcatcctg	540
gggatcgctg	tgctggtagg	cagcttcgtc	gcgaacagg	tctaccagag	cggccacctg	600
gccgagtgc	tactggtcac	gacgtatctg	ctgatcgcaa	ccgccgcagc	gctggccgtg	660
ccgttgccgc	gcgggtgcaa	ctcgttgggg	gcgccacaag	ttgccggcgc	cgctacggcc	720
gtgctgtttt	tgaccttgat	gacgcggggc	ggccctcgga	agcgtcatga	gttggcgctg	780
tttgccgtga	tcaccgctat	cgcggtcatc	gcggccgcgc	ctgccttcgg	ctatggatac	840
caggactggg	tccccgcggg	ggggatcgca	ttcgggctgt	tcattgtgac	gaatgcggcc	900
aagctgaccg	tcgcggtcgc	gcggatcgcg	ctgccgcgca	ttccggtacc	cggcgaaacc	960
gtggacaacg	aggagttgct	cgatcccgtc	gcgaccccg	aggctaccag	cgaagaaacc	1020
ccgacctggc	aggccatcat	cgcgtcgggt	cccgcgtccg	cggctccggc	caccgagcgc	1080
agcaaaactg	ccaagcaact	tctgatcgga	tacgtcacgt	cgggcaccct	gattctggct	1140
gccggtgcca	tcgcggtcgt	ggtgcgcggg	cacttctttg	tacacagcct	ggtggtcgcg	1200
ggtttgatca	cgaccgtctg	cggatttcgc	tcgcggcttt	acgccgagcg	ctggtgtgcg	1260
tgggcgttgc	tggcggcgac	ggtcgcgatt	ccgacgggtc	tgacggccaa	actcatcatc	1320
tggtagccgc	actatgcctg	gctgttgttg	agcgtctacc	tcacggtagc	cctggttgcg	1380
ctcgtggtgg	tcgggtcgat	ggctcacgtc	cggcgcgttt	caccggtcgt	aaaacgaact	1440
ctggaattga	tcgacggcgc	catgatcgct	gccatcatte	ccatgctgct	gtggatcacc	1500
ggggtgtacg	acacggtccg	caatatccgg	ttctga			1536

&lt;210&gt; 21

&lt;211&gt; 840

&lt;212&gt; DNA

&lt;213&gt; Mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; DNA sequence Rv3878

&lt;400&gt; 21

atggctgaac	cggttgccgt	cgatcccacc	ggcttgagcg	cagcggccgc	gaaattggcc	60
ggcctcgttt	ttccgcagcc	tccggcgccg	atcgcggtca	gcggaacgga	ttcggtggta	120
gcagcaatca	acgagaccat	gccaagcatc	gaatcgctgg	tcagtgacgg	gctgcccggc	180
gtgaaagccg	ccctgactcg	aacagcatcc	aacatgaacg	cggcggcgga	cgtctatgcg	240
aagaccgatc	agtcactggg	aaccagtttg	agccagtatg	cattcggtc	gtcgggcgaa	300
ggcctggctg	gcgtcgccctc	ggtcgggtgg	cagccaagtc	aggctacca	gctgctgagc	360
acacccggtg	cacaggtcac	gaccagctc	ggcgagacgg	ccgtgagct	ggcaccctgt	420
gttgttgca	cggtagccga	actcgttcag	ctggctccgc	acgccgttca	gatgtcgcaa	480
aacgcatacc	ccatcgctca	gacgatcagt	caaaccgccc	aacaggccgc	ccagagcgcg	540
cagggcgga	gcggcccaat	gcccgcacag	cttgccagcg	ctgaaaaacc	ggccaccgag	600
caagcggagc	cgggtccacga	agtgacaaac	gacgatcagg	gcgaccaggg	cgacgtgcag	660
ccggccgagg	tcgttgccgc	ggcacgtgac	gaaggcgccg	gcgcatacc	gggccagcag	720
cccggcgggg	gcgttcccgc	gcaagccatg	gataccggag	ccggtgccc	cccagcgcg	780
agtcgcgtgg	cggcccccgt	cgatccgctg	actccggcac	cctcaacaac	cacaacgttg	840

&lt;210&gt; 22

&lt;211&gt; 2187

&lt;212&gt; DNA

## &lt;213&gt; Mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; DNA sequence Rv3879c

&lt;400&gt; 22

```

atgagtatta ccaggccgac gggcagctat gccagacaga tgctggatcc gggcggctgg 60
gtggaagccg atgaagacac tttctatgac cgggccagg aatatagcca ggttttgcaa 120
agggtcaccg atgtattgga cacctgccgc cagcagaaag gccacgtctt cgaaggcggc 180
ctatggctcg gggcgccgc caatgctgcc aacggcgccc tgggtgcaaa catcaatcaa 240
ttgatgacgc tgcaggatta tctcgccacg gtgattacct ggcacaggca tattgccggg 300
ttgattgagc aagctaaatc cgatatcggc aataatgtgg atggcgctca acgggagatc 360
gatatcctgg agaatgaccc tagcctggat gctgatgagc gccataccgc catcaattca 420
ttggtcacgg cgacgcattg ggccaatgtc agtctggctg ccgagaccgc tgagcgggtg 480
ctggaatcca agaattgga acctccgaag aacgcactcg aggatttgct tcagcagaag 540
tcggcgccac cccagacgt gcctaccctg gtcgtgccat ccccgggcac accgggcaca 600
ccgggaaccc cgatcacccc gggaaccccg atcaccccg gaacccaat cacaccatc 660
ccgggagcgc cggtaaactc gatcacacca acgcccggca ctcccgctac gccggtgacc 720
ccgggcaagc cggtcacccc ggtgaccccg gtcaaaccgg gcacaccagg cgagccaacc 780
ccgatcacgc cggtcacccc ccggtcgcc ccggccacac cggcaacccc ggccacgccc 840
gttaccacag ctcccgctcc acaccgcag ccggtccgg caccggcgcc atcgctggg 900
ccccagccgg ttacaccggc cactcccggg ccgtctgggtc cagcaacacc gggcacccca 960
gggggcgagc cggcgcgcga cgtcaaacc cggcggttg cgagcaacc tgggtgtgcc 1020
ggccagcatg cgggcggggg gacgcagtcg ggcctgccc atgcggacga atccgcgcg 1080
tcgggtgacg cggtcgcggc gtccggtgtc ccggcgccac gggcggcggc cgccgcgcg 1140
agcgttaccg ccgtgggagc ggcgcgcgt tcgagcgtgg gtacggccgc ggcctcgggc 1200
gcggggtcgc atgctgccac tgggcgggcg ccggtggcta cctcggacaa ggcggcggca 1260
ccgagcacgc gggcgccctc ggcgcggacg gcacctctg cccgcccgc gtcgaccgat 1320
cacatcgaca aaccgatcg cagcgagtct gcagatgac gtacgccggg ttcgatgatc 1380
ccggtgtcgc cggtcgggc ggcacgcgac gccgccactg cagctgccag cgcccgccag 1440
cgtggccgcg gtgatgcgct gcggttggcg cgacgcacg cggcgccgct caacgcgtcc 1500
gacaacaacg cgggcgacta cgggttcttc tggatcaccc cggtgaccac cgacggttcc 1560
atcgtcgtgg ccaacagcta tgggttgccc tacatacccg acgggatgga attgccgaat 1620
aaggtgtact tggccagcgc ggatcacgca atcccggttg acgaaattgc acgctgtgcc 1680
acctaccggg ttttgccgt gcaagcctgg gcggctttcc acgacatgac gctgcgggcg 1740
gtgatcggtg ccgcggagca gttggccagt tcggatcccg gtgtggccaa gattgtgctg 1800
gagccagatg acattccgga gagcggcaaa atgacgggccc ggtcgcggct ggaggtcgtc 1860
gacccctcgg cggcggctca gctggccgac actaccgatc agcgtttgct cgacttggtg 1920
ccgcggcgcc cggtggatgt caatccaccg ggcatgagc ggcacatgct gtggttcgag 1980
ctgatgaagc ccatgaccag caccgctacc ggccgcgagg ccgctcatct gcgggcttcc 2040
cgggcctacg ctgcccactc acaggagatt gccctgcacc aagcgcacac tgcgactgac 2100
gcggccgtcc agcgtgtggc cgtcgcgac tggctgtact ggcaatacgt caccgggttg 2160
ctcgaaccgg ccctggccgc cgcattgc

```

&lt;210&gt; 23

&lt;211&gt; 345

&lt;212&gt; DNA

&lt;213&gt; Mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; DNA sequence Rv3880c

&lt;400&gt; 23

```

gtgagcatgg acgaattgga cccgcatgtc gcccgggcgt tgacgctggc ggcgcggttt 60
cagtcggccc tagacgggac gctcaatcag atgaacaacg gatccttcg cgccaccgac 120
gaagccgaga ccgtcgaagt gacgatcaat gggcaccagt ggctcaccgg cctgcgcac 180
gaagatgggt tgctgaagaa gctgggtgcc gaggcgggtg ctcagcgggt caacgagggc 240
ctgcacaatg cgcaggccgc ggcgtccgcg tataacgacg cggcgggcga gcagctgacc 300
gctgcgttat cggccatgtc ccgcgcgatg aacgaaggaa tggcc

```

<210> 24  
 <211> 1380  
 <212> DNA  
 <213> Mycobacterium tuberculosis

<220>  
 <223> DNA sequence Rv3881c

<400> 24  
 atgacgcagt cgcagaccgt gacggtggat cagcaagaga ttttgaacag ggccaacgag 60  
 gtggaggccc cgatggcgga cccaccgact gatgtcccca tcacaccgtg cgaactcacg 120  
 gcggctaaaa acgccgccc acagctggta ttgtccgccg acaacatgcy ggaataacctg 180  
 gcggccggtg ccaaagagcg gcagcgtctg gcgacctgc tgcgcaacgc ggccaaggcg 240  
 tatggcgagg ttgatgagga ggctgcgacc gcgctggaca acgacggcga aggaactgtg 300  
 caggcagaat cggccggggc cgtcggaggg gacagttcgg ccgaactaac cgatacgccg 360  
 agggtggcca cggccggtga acccaacttc atggatctca aagaagcggc aaggaagctc 420  
 gaaacggggc accaaggcgc atcgtctcgc cactttgcgg atgggtggaa cactttcaac 480  
 ctgacgctgc aaggcgacgt caagcggttc cgggggtttg acaactggga aggcgatgcy 540  
 gctaccgctt gcgaggcttc gctcgatcaa caacggcaat ggatactcca catggccaaa 600  
 ttgagcgctg cgatggccaa gcaggctcaa tatgtcgcgc agctgcacgt gtgggctagg 660  
 cgggaacatc cgacttatga agacatagtc gggctcgaac ggctttacgc ggaaaaccct 720  
 tcggcccgcg accaaattct cccggtgtac gcggagtatc agcagaggtc ggagaagggtg 780  
 ctgaccgaat acaacaacaa ggcagccctg gaaccggtaa acccgccgaa gcctcccccc 840  
 gccatcaaga tcgaccgcgc cccgcctccg caagagcagg gattgatccc tggcttcctg 900  
 atgcccgcgt ctgacggctc cgggtgtgact cccggtaccg ggatgccagc cgcaccgatg 960  
 gttccgccta ccggtatgcc ggggtggggc ctcccggtg acacggcggc gcagctgacg 1020  
 tcggctgggg ggggaagcgc agcgtgtcgc ggcgacgtgg cgggtcaaagc ggcacgctc 1080  
 ggtggcggtg gaggcggcgc ggtgccgtcg gcgccgttgg gatccgcgat cgggggcgc 1140  
 gaatcggtgc ggcccgctgg cgtcgggtgac attgccggct taggccaggg aaggggcggc 1200  
 ggcgcgcccg cgtcgggcgc cgggtggcatg ggaatgccga tgggtgccgc gcatcagggg 1260  
 caagggggcg ccaagtccaa gggttctcag caggaagacg aggcgctcta caccgaggat 1320  
 cgggcatgga ccgaggccgt cattggtaac cgtcggcgcc aggacagtaa ggagtcgaag 1380

<210> 25  
 <211> 1386  
 <212> DNA  
 <213> Mycobacterium tuberculosis

<220>  
 <223> DNA sequence Rv3882c

<400> 25  
 atgagaaatc ctttaggggt gcggttcagc accgggcacg ccttgcttgc ctccgcgttg 60  
 gccccgccat gcatcatcgc attcttggag acgcgctact ggtgggcggg gattgcgctg 120  
 gcctcgtttg gcgtcatcgt ggccacggtc actttctacg gccgccgat caccggctgg 180  
 gtggcgggcg tgtacgcgtg gttgcggcgc cgccgacggc ccccgattc ctcgtcagaa 240  
 cctgtggtcg gggccaccgt gaagccagga gatcacgttg cgggtgcgtg gcaaggcgag 300  
 tttctggtcg ccgtaatcga gctcattccc cgaccattca cgccgacggg catcgctcac 360  
 gggcaagccc acaccgacga catgctggac accggactgg tggaggagct cctgtcgggtg 420  
 cactgtcccc acttgaggc cgatatcgtc tcagccgggt accgcgtcgc caataccgca 480  
 gcgcggagc ggtgtgagtc gtatcagcag gtgatcgga cagaccggc gccggcgaac 540  
 cgccggacct ggatcggtg gcgcgccgac ccggaacgca cccgcaaadc ggcgcagcgc 600  
 cgcgatgaag gcgtcgcagg actggcccgc tatttgggtg cgtccgcgac gcgcattgcc 660  
 gatcgactgg ctagccatgg tgtcgacgcg gtgtgtggcc gcagcttcga tgactacgac 720  
 caccgccacc acatcggtct tgtcggggag aaatggctga tgatcaaggg gcgcgatgcc 780  
 tacactgccg cctacgcggc gcccgagggt ccggatgtat ggtggctcgg gcgcgcggac 840  
 cacaccatca ccagagtcgc ggtcgcgcgc gggatggccc cgcagtcac ggtgttgctg 900  
 accacggcgg acaagcccaa gacaccagc ggcttcgccc gcctatttgg cgggcagcgg 960  
 cccgcgctgc aaggccagca tctggtggcc aaccgccact gccagctgcc gatcgggtca 1020

```

gctgggggtac tgggtcggcga gacgggtgaac cgatgcccg tctacatgcc cttcgacgat 1080
gtcgacatcg ccctcaacct ggggtgacgct cagacattca cccagttcgt ggtgcggtgcg 1140
gcggcggcag gtgcgatggg cacagtcggg ccacagttcg aggaatttgc ccggttgatc 1200
ggcgcacaca tcgggcagga ggtaaagggtg gcgtggccga atgcgacgac ctatctcggc 1260
ccgcatcccg gtattgaccg ggtgattctg cggcacaatg tgatcggtac cccgcggcat 1320
cggcagctgc cgattcgccg ggtttcccca cccgaggaaa gccgctacca gatggcgctg 1380
ccgaag
1386

```

&lt;210&gt; 26

&lt;211&gt; 1338

&lt;212&gt; DNA

&lt;213&gt; Mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; DNA sequence Rv3883c

&lt;400&gt; 26

```

gtgcaccgta tctttctgat cacgggtggcg ctggcggttg tcaccgcgtc gcccgcatcg 60
gccatcacgc caccgccgat cgatccgggc gcgttgccgc ccgacgtgac gggcccgat 120
cagcctaccg aacagcgcggt tttgtgcgcg tcgcccacca cgctgccggg gtccgggttc 180
cacgatccgc cgtggagcaa cacgtatctg ggctggcccg atgccacaaa gtccgcgacc 240
ggggccgggg tgacgggtggc ggtgatcgac accggtgtcg acgcttcgcc acgggtccc 300
gcggaacctg gcggcgattt cgtcgaccag gccggtaacg gctgtctga ctgtgatgcc 360
catgggactc tcacagcatc catcatcgcg ggccggcccg cggccaccga cgggttcgtc 420
ggcgctcgcg ccgacgctcg actgctctcg ctacgtcaga cgtctgaggc cttcgaaccg 480
gtcggctcac aagccaaccc gaatgacccc aacgccaccc cggccgcggg ttccatccgc 540
agtcttgccc gcgcggtggg gacgcgccgc aacctcggcg tgggtgtgat caacatcagt 600
gaagccgctt gctacaaggt gagcaggccg atcgatgaaa cctcactggg tgcattccatc 660
gactatgcgg tcaacgtcaa aggcgtgggt gtgggtggtc cggccggcaa caccggtggc 720
gattgcgtac agaattccggc gccggacccg tccacacccg gcgacccacg cggctggaac 780
aatgtgcaga ccgttgtcac cccggcggtg tacgcaccgc tgggtgttaag cgtcggcggt 840
atcggccaga ccgggatgcc cagctcgttc tcgatgcacg gaccgtgggt ggacgtggcc 900
gcgcccgcag aaaacatcgt cgcgctcggc gacaccggtg aaccggtgaa tgcgctgcaa 960
ggccgggagg ggccgggtacc catcgccggc acctcgtttg ccgcggcata tgtgtcgggt 1020
ctggcgggccc tgcttcggca gcgggtcccc gacctgacgc cggcgagat catccaccgg 1080
atcaccgcca ccgcgagaca ccccgggggc ggggtcgacg acctggtcgg cgcggcgctc 1140
atcgatgcgg tggccgcgct gacgtgggac attccgcccg gccctgcttc ggcgccatac 1200
aacgtcagac gatttcacc cccggtgggt gagccgggtc ccgatcgctc cccgattacg 1260
gctgtggcgt tgggtggcgt cggccttacg ttggccctgg gcctgggcgc gctggctaga 1320
cgggcgctga gccgcga
1338

```

&lt;210&gt; 27

&lt;211&gt; 1857

&lt;212&gt; DNA

&lt;213&gt; Mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; DNA sequence Rv3884c

&lt;400&gt; 27

```

atgtcgagaa tgggtggacac gatgggtgat ttactcactg cgcgccggca tttcgatcgg 60
gcgatgacga tcaagaatgg ccagggatgc gtggcgcggt tgctgagtt tgtggctgcc 120
accgaggccg atcgcgtgat ggccgacgcg tggttgggtc gtatcgctc cggtgaccgc 180
gatctggcct cgcttaagca gctcaacgcc catagcgagt ggctgcaccg cgagaccacg 240
cggatcggcc ggacgttggc cgctgagggt cagctgggac catccatcgg gatcacggtg 300
accgacgcat ctcagggtggg gctggcgctg tcgtcggcgt tgacgatcgc gggggagtat 360
gcgaaggccg atgccctgtt agcaaacccg gagctattgg attcgtggcg caactaccag 420
tggcatcagc tggctcgggc gttcctgatg tacgtcacgc agcgatggcc cgacgtgttg 480
tcgacggccg ccgaggatct gccgccacag gcgatcgta tgccggcggt gaccgcgtcg 540

```



atttgtgcgc	tggcagccca	cgccgccgcc	catctcgggc	aggggcgagt	ggccctggac	600
tggctggacc	gggtggacgt	gatcggacac	agcaggatcat	cggagcgggt	cgccgccgac	660
gtgctcaccg	cggcgatcgg	accggccgat	attccgctgc	tggtcgccga	cttggcgtat	720
gtgcggggga	tgggtgtaccg	gcaactgcac	gaggaggaca	aggcccagat	ctggctgtcg	780
aaggccacca	tcaacggggg	gtcaccgcac	gccgccaaaag	aagccctggc	ggaccggaac	840
ctgcgcttga	ttgtttaccga	tgaacgaacc	atcgccagcc	gctccgaccg	ttgggatgct	900
tcgacggcga	aaagccgcga	ccagctcgat	gacgacaatg	cagcgcagcg	gcgcggcgag	960
ctgctagccg	agggccggga	actgctggcc	aaacagggtg	gcctggcgcc	ggtcaagcaa	1020
gcggtatcgg	cgctggaaga	ccaactcgag	gtgcgcacga	tgcgcctaga	gcacggccta	1080
ccggtggagg	ggcagaccaa	ccacatgttg	ctggtggggc	caccaggcac	aggtaagaca	1140
accaccgctg	aagcgctcgg	caagatctac	gccggcatgg	ggatcggtcg	tcaccccgaa	1200
attcgagaag	ttcgccgatc	ggacttctgt	gggcactaca	tcggggagtc	aggacccaag	1260
acgaacgagc	tgatcgaaaa	gtcactcggg	cgaatcattt	tcattggacga	gttctactcg	1320
ctgatcgaac	gtcatcaaga	cggaacaccg	gacatgatcg	gcattggaggc	ggtcaatcaa	1380
ctcctgggtc	aattggaaaac	acaccgattc	gacttctgtt	tcattcgggc	cggctatgag	1440
gatcagggtg	atgaattcct	caccgtgaac	ccgggtttgg	ctggccgggt	caaccgaaaag	1500
ctgcggttcg	agtctttatc	gccgggtggag	atcgctcgaga	ttggacaccg	ctacgctaca	1560
ccgcgcgcca	gccagctcga	tgacgcccga	ccggagggtat	tcctcgacgc	ggtcaccacc	1620
atccgtaact	acaccacccc	tagtgggcag	cacgggtatcg	acgctatgca	aaacggtcgg	1680
ttcgcccgcga	acgtgatcga	acgcgccgaa	gggttccggg	acaccgggtg	ggttgcgcaa	1740
aaacgtgcgg	gccaaccggg	atcggttcag	gatctgcaga	tcattaccgc	caccgacatc	1800
gatgccgcga	tacgcagcgt	gtgctcagac	aaccgagaca	tggccgcgat	cgtttgg	1857

&lt;210&gt; 28

&lt;211&gt; 1611

&lt;212&gt; DNA

&lt;213&gt; Mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; DNA sequence Rv3885c

&lt;400&gt; 28

ttgacgtcca	agctgaccgg	gttcagcccc	cgcagtgcga	ggcgggtcgc	cggggtgtgg	60
acagtgttcg	tgctcgcgtc	ggcgggctgg	gcgctggggc	gccagctagg	tgcggtcatg	120
gctgtcgtgg	tcggcgctgc	cttgggtgtc	gtacagtggg	ggggtcagcc	ggcgtggctg	180
tgggcggtac	tggggctgcg	gggtcggcgt	cccgtcaaag	ggaatgacc	aattaccttg	240
gccaacaacc	gatccggggg	tggcgctccg	gtgcaagacg	gtgtcgcggg	ggtggcgggtg	300
caacttctcg	gccgagcgca	ccggggcgact	acggtcaccg	ggtcgggtgac	cgtagaaagc	360
gacaacgtga	ttgacgtcgt	tgagctcgcg	ccgttgcgtg	gccacccgct	ggacctggaa	420
ctcgattcaa	tcagcgtcgt	caccttcggc	tcgcgaaccg	gcaccgtcgg	cgattaccgg	480
cgggtgtatg	acgcggagat	cggtacgccg	ccgtatgccg	ggcggcgcca	aacgtggctg	540
atcatgcggc	ttccggtgat	cggcaacacc	caagctttac	gctggcgtag	cagcgttggg	600
gccgctgccca	tttcgggtcgc	ccaacgcggt	gccagctccc	tgcgctgtca	gggcttgcgc	660
gccaacttgg	ccaccgcaac	agacttggct	gagcttgcgc	gccggctggg	gtcggacgcg	720
gtagccggga	gtgcgcagcg	ctggaaaagt	atccgcgggtg	aagccgggtg	gatgacgacg	780
tatgcgtacc	cggctgaggg	gatttcgtcg	cgggttctct	cgcaagcctg	gacgctgcgt	840
gccgatgagg	tcattccagaa	cgtaacgggtg	tatccggacg	cgacgtgcac	cgcgaccatc	900
accgtgcgca	caccgacgcc	ggcgccctacc	ccgcccagtg	tgatcttgcg	tcggctcaat	960
ggtgagcaag	ccgcgcgggc	tgcggccaac	atgtgcgggc	cacgtccaca	cctacgcgga	1020
cagcggcgct	gcccgtttgc	ggcgcagcta	gtcaccgaga	tcggaccgtc	gggggtgttg	1080
attggcaagc	tgagcaacgg	ggaccggctg	atgattcccc	ttaccgacgc	cggtagagctg	1140
tcgcgcgtct	tcgtggccgc	ggacgacaag	atcgccaaga	ggatcggtgat	tcgcgtcgtc	1200
ggtgccgggtg	agcgggtgtg	tgtgcacact	cgcgaccaag	agcgttgggc	cagcgtacgc	1260
atgccgcagt	tgagcatcgt	cggcacacca	cggcccgcgc	cgcgaccac	tgtcggcgctc	1320
gtggagtacg	tgcggcgggc	caagaacggc	gatgacggca	aatctgaagg	cagcgggtgtc	1380
gatgtcgcga	tttcgcccac	gccacggcca	gccagtgtca	tcaccattgc	ccgaccgggt	1440
acctcgctgt	ccgagagcga	tcggcatggc	ttcgagggtga	ccatcgaaac	aatcgatcgg	1500
gcaacgggtga	aagtcgggtg	ggcaggacaa	aactggctgg	ttgagatgga	aatgttccgt	1560
gcggagaacc	gctatgtcag	ccttgagccc	gtcacgatgt	cgataggccg	g	1611

<210> 29  
 <211> 620  
 <212> DNA  
 <213> Mycobacterium tuberculosis

<220>  
 <223> CFP-10 + ESAT-6

<400> 29  
 atggcagaga tgaagaccga tgccgctacc ctccgcgcagg aggcaggtaa tttcgagcgg 60  
 atctccggcg acctgaaaac ccagatcgac caggtggagt cgacggcagg ttcggtgcag 120  
 ggccagtggc gccgcgcggc ggggacggcc gccaggccg cgggtggtgcg cttccaagaa 180  
 gcagccaata agcagaagca ggaactcgac gagatctcga cgaatattcg tcaggccggc 240  
 gtccaatact cgagggccga cgaggagcag cagcaggcgc tgcctcgcga aatgggcttc 300  
 tgacccgcta atacgaaaag aaacggagca aaaacatgac agagcagcag tggaatttcg 360  
 cgggtatcga ggccgcggca agcgcaatcc agggaaatgt cacgtccatt cattccctcc 420  
 ttgacgaggg gaagcagtcc ctgaccaagc tcgcagcggc ctggggcggg agcggttcgg 480  
 aggcgtacca ggggtgtccag caaaaatggg acgccacggc taccgagctg aacaacgcgc 540  
 tgcagaacct ggccgcggacg atcagcgaag ccggtcaggc aatggcttcg accgaaggca 600  
 acgtcactgg gatgttcgca 620

<210> 30  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer SP6-BAC1

<400> 30  
 agttagctca ctcattagcg a 21

<210> 31  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer T7-BAC1

<400> 31  
 ggatgtgctg caaggcgatt a 21

<210> 32  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer esat-6F

<400> 32  
 gtcacgtcca ttcattccct 20

<210> 33  
 <211> 19

<212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer esat-6R

<400> 33  
 atcccagtg cgttgcctt

19

<210> 34  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer RD1<sup>mic</sup>  
 flanking region F

<400> 34  
 gcagtgcaaa ggtgcagata

20

<210> 35  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer RD1<sup>mic</sup>  
 flanking region R

<400> 35  
 gattgagaca cttgccacga

20

<210> 36  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer  
 plcA.int.F

<400> 36  
 caagttgggt ctggtcgaat

20

<210> 37  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer  
 plcA.int.R

<400> 37  
 gctaccaag gtctcctggt

20

<210> 38  
 <211> 153  
 <212> DNA  
 <213> Mycobacterium tuberculosis  
 <220>  
 <223> Sequences at the junction RD1<sup>mic</sup>

<400> 38  
 caagacgagg ttgtaaaacc tcgacgcagg atcggcgatg aaatgccagt cggcgtcgct 60  
 gagcgcgcg cgcgccgagt cccattttgt cgctgatttg tttgaacagc gacgaaccgg 120  
 tgttgaaaat gtcgcctggg tcggggattc cct 153

<210> 39  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer RD5<sup>mic</sup>  
 flanking region F

<400> 39  
 gaatgccgac gtcatatcg 19

<210> 40  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer RD5<sup>mic</sup>  
 flanking region R

<400> 40  
 cggccactga gttcgattat 20

<210> 41  
 <211> 152  
 <212> DNA  
 <213> Mycobacterium tuberculosis

<220>  
 <223> Sequence at the junction RD5<sup>mic</sup>

<400> 41  
 cctcgatgaa ccacctgaca tgaccccatc ctttccaaga actggagtct cgggacatgc 60  
 cggggcggtt cactgcccga ggtgtcctgg gtcgttccgt tgaccgtcga gtccgaacat 120  
 ccgtcattcc cgggtggcagt cgggtgcggtg ac 152

<210> 42  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer MiD1  
flanking region F

<400> 42  
cagccaacac caagtagacg 20

<210> 43  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer MiD1  
flanking region R

<400> 43  
tctacctgca gtcgcttg 20

<210> 44  
<211> 123  
<212> DNA  
<213> Mycobacterium tuberculosis

<220>  
<223> Sequence at the junction MiD1

<400> 44  
cacctgacat gaccccatcc tttccaagaa ctggagtctc cggacatgcc ggggcgggttc 60  
agggacattc atgtccatct tctggcagat cagcagatcg cttgttctca gtgcaggtga 120  
gtc 123

<210> 45  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer MiD2  
flanking region R

<400> 45  
gtccatcgag gatgtcgagt 20

<210> 46  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer MiD2  
flanking region L

<400> 46  
ctaggccatt ccgttgtctg 20

<210> 47

<211> 151  
 <212> DNA  
 <213> Mycobacterium tuberculosis

<220>  
 <223> Sequence at the junction MiD2

<400> 47  
 gctgcctact acgctcaacg ccagagacca gccgccggct gaggtctcag atcagagagt 60  
 ctccggactc accggggcgg ttcataaagg cttcgagacc ggacgggctg taggttcctc 120  
 aactgtgtgg cggatggtct gaggacttaa c 151

<210> 48  
 <211> 15  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer MiD3  
 flanking region R

<400> 48  
 ggcgacgcca tttcc 15

<210> 49  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer MiD3  
 flanking region L

<400> 49  
 aactgtcggg cttgtcttt 19

<210> 50  
 <211> 181  
 <212> DNA  
 <213> Mycobacterium tuberculosis

<220>  
 <223> Sequence at the junction MiD3

<400> 50  
 tggcgccggc acctccgttg ccaccgttgc cgccgctggt gggcgcggtg ccgttcgccc 60  
 cggccgaacc gtccagggcc gggttcgccc tcagccgcta aacacgccga ccaagatcaa 120  
 cgagctacct gcccggtcaa ggttgaagag ccccatatc agcaagggcc cgggtgtcggc 180  
 g 181

<210> 51  
 <211> 108  
 <212> PRT  
 <213> Mycobacterium tuberculosis

<220>  
 <223> RV3861 - hypothetical protein



&lt;400&gt; 51

Val Thr Trp Leu Ala Asp Pro Val Gly Asn Ser Arg Ile Ala Arg Ala  
 1 5 10 15

Gln Ala Cys Lys Thr Ser Ile Ser Ala Pro Ile Val Glu Ser Trp Arg  
 20 25 30

Ala Gln Arg Gly Ala Gln Cys Gly Gln Arg Glu Lys Ser Cys Arg Cys  
 35 40 45

Ser Arg Ala Val His Ile Gln Gly Ile Ser Pro Pro Leu Phe Arg Arg  
 50 55 60

Pro Leu Glu Pro Ala Val Gln Ala Ala Val Ala Ser Cys Arg Leu Gly  
 65 70 75 80

Arg His Pro Val Val Ala His Arg Val Thr Val Ala Leu Gly Gln Gly  
 85 90 95

Ser Gln Leu Ala Gln Arg Glu Cys Pro Arg Pro Ala  
 100 105

&lt;210&gt; 52

&lt;211&gt; 116

&lt;212&gt; PRT

&lt;213&gt; Mycobacterium tuberculosis

&lt;220&gt;

<223> WHIB6 - Possible transcriptional regulatory  
 protein WHIB-like WHIB6

&lt;400&gt; 52

Met Arg Tyr Ala Phe Ala Ala Glu Ala Thr Thr Cys Asn Ala Phe Trp  
 1 5 10 15

Arg Asn Val Asp Met Thr Val Thr Ala Leu Tyr Glu Val Pro Leu Gly  
 20 25 30

Val Cys Thr Gln Asp Pro Asp Arg Trp Thr Thr Thr Pro Asp Asp Glu  
 35 40 45

Ala Lys Thr Leu Cys Arg Ala Cys Pro Arg Arg Trp Leu Cys Ala Arg  
 50 55 60

Asp Ala Val Glu Ser Ala Gly Ala Glu Gly Leu Trp Ala Gly Val Val  
 65 70 75 80

Ile Pro Glu Ser Gly Arg Ala Arg Ala Phe Ala Leu Gly Gln Leu Arg  
 85 90 95

Ser Leu Ala Glu Arg Asn Gly Tyr Pro Val Arg Asp His Arg Val Ser  
 100 105 110

Ala Gln Ser Ala  
 115

&lt;210&gt; 53

&lt;211&gt; 392

&lt;212&gt; PRT

<213> Mycobacterium tuberculosis

<220>

<223> Rv3863 - hypothetical alanine rich protein

<400> 53

Met Ala Gly Glu Arg Lys Val Cys Pro Pro Ser Arg Leu Val Pro Ala  
1 5 10 15

Asn Lys Gly Ser Thr Gln Met Ser Lys Ala Gly Ser Thr Val Gly Pro  
20 25 30

Ala Pro Leu Val Ala Cys Ser Gly Gly Thr Ser Asp Val Ile Glu Pro  
35 40 45

Arg Arg Gly Val Ala Ile Ile Gly His Ser Cys Arg Val Gly Thr Gln  
50 55 60

Ile Asp Asp Ser Arg Ile Ser Gln Thr His Leu Arg Ala Val Ser Asp  
65 70 75 80

Asp Gly Arg Trp Arg Ile Val Gly Asn Ile Pro Arg Gly Met Phe Val  
85 90 95

Gly Gly Arg Arg Gly Ser Ser Val Thr Val Ser Asp Lys Thr Leu Ile  
100 105 110

Arg Phe Gly Asp Pro Pro Gly Gly Lys Ala Leu Thr Phe Glu Val Val  
115 120 125

Arg Pro Ser Asp Ser Ala Ala Gln His Gly Arg Val Gln Pro Ser Ala  
130 135 140

Asp Leu Ser Asp Asp Pro Ala His Asn Ala Ala Pro Val Ala Pro Asp  
145 150 155 160

Pro Gly Val Val Arg Ala Gly Ala Ala Ala Ala Arg Arg Arg Glu  
165 170 175

Leu Asp Ile Ser Gln Arg Ser Leu Ala Ala Asp Gly Ile Ile Asn Ala  
180 185 190

Gly Ala Leu Ile Ala Phe Glu Lys Gly Arg Ser Trp Pro Arg Glu Arg  
195 200 205

Thr Arg Ala Lys Leu Glu Glu Val Leu Gln Trp Pro Ala Gly Thr Ile  
210 215 220

Ala Arg Ile Arg Arg Gly Glu Pro Thr Glu Pro Ala Thr Asn Pro Asp  
225 230 235 240

Ala Ser Pro Gly Leu Arg Pro Ala Asp Gly Pro Ala Ser Leu Ile Ala  
245 250 255

Gln Ala Val Thr Ala Ala Val Asp Gly Cys Ser Leu Ala Ile Ala Ala  
260 265 270

Leu Pro Ala Thr Glu Asp Pro Glu Phe Thr Glu Arg Ala Ala Pro Ile  
275 280 285

Leu Ala Asp Leu Arg Gln Leu Glu Ala Ile Ala Val Gln Ala Thr Arg

290

295

300

Ile Ser Arg Ile Thr Pro Glu Leu Ile Lys Ala Leu Gly Ala Val Arg  
305 310 315 320

Arg His His Asp Glu Leu Met Arg Leu Gly Ala Thr Ala Pro Gly Ala  
325 330 335

Thr Leu Ala Gln Arg Leu Tyr Ala Ala Arg Arg Arg Ala Asn Leu Ser  
340 345 350

Thr Leu Glu Thr Ala Gln Ala Ala Gly Val Ala Glu Glu Met Ile Val  
355 360 365

Gly Ala Glu Ala Glu Glu Glu Leu Pro Ala Glu Ala Thr Glu Ala Ile  
370 375 380

Glu Ala Leu Ile Arg Gln Ile Asn  
385 390

<210> 54

<211> 402

<212> PRT

<213> Mycobacterium tuberculosis

<220>

<223> Rv3864 - conserved hypothetical protein

<400> 54

Met Ala Ser Gly Ser Gly Leu Cys Lys Thr Thr Ser Asn Phe Ile Trp  
1 5 10 15

Gly Gln Leu Leu Leu Gly Glu Gly Ile Pro Asp Pro Gly Asp Ile  
20 25 30

Phe Asn Thr Gly Ser Ser Leu Phe Lys Gln Ile Ser Asp Lys Met Gly  
35 40 45

Leu Ala Ile Pro Gly Thr Asn Trp Ile Gly Gln Ala Ala Glu Ala Tyr  
50 55 60

Leu Asn Gln Asn Ile Ala Gln Gln Leu Arg Ala Gln Val Met Gly Asp  
65 70 75 80

Leu Asp Lys Leu Thr Gly Asn Met Ile Ser Asn Gln Ala Lys Tyr Val  
85 90 95

Ser Asp Thr Arg Asp Val Leu Arg Ala Met Lys Lys Met Ile Asp Gly  
100 105 110

Val Tyr Lys Val Cys Lys Gly Leu Glu Lys Ile Pro Leu Leu Gly His  
115 120 125

Leu Trp Ser Trp Glu Leu Ala Ile Pro Met Ser Gly Ile Ala Met Ala  
130 135 140

Val Val Gly Gly Ala Leu Leu Tyr Leu Thr Ile Met Thr Leu Met Asn  
145 150 155 160

Ala Thr Asn Leu Arg Gly Ile Leu Gly Arg Leu Ile Glu Met Leu Thr

165					170					175					
Thr	Leu	Pro	Lys	Phe	Pro	Gly	Leu	Pro	Gly	Leu	Pro	Ser	Leu	Pro	Asp
			180					185					190		
Ile	Ile	Asp	Gly	Leu	Trp	Pro	Pro	Lys	Leu	Pro	Asp	Ile	Pro	Ile	Pro
		195					200					205			
Gly	Leu	Pro	Asp	Ile	Pro	Gly	Leu	Pro	Asp	Phe	Lys	Trp	Pro	Pro	Thr
	210					215					220				
Pro	Gly	Ser	Pro	Leu	Phe	Pro	Asp	Leu	Pro	Ser	Phe	Pro	Gly	Phe	Pro
225						230					235				240
Gly	Phe	Pro	Glu	Phe	Pro	Ala	Ile	Pro	Gly	Phe	Pro	Ala	Leu	Pro	Gly
				245					250					255	
Leu	Pro	Ser	Ile	Pro	Asn	Leu	Phe	Pro	Gly	Leu	Pro	Gly	Leu	Gly	Asp
			260					265					270		
Leu	Leu	Pro	Gly	Val	Gly	Asp	Leu	Gly	Lys	Leu	Pro	Thr	Trp	Thr	Glu
		275					280					285			
Leu	Ala	Ala	Leu	Pro	Asp	Phe	Leu	Gly	Gly	Phe	Ala	Gly	Leu	Pro	Ser
	290					295					300				
Leu	Gly	Phe	Gly	Asn	Leu	Leu	Ser	Phe	Ala	Ser	Leu	Pro	Thr	Val	Gly
305				310					315					320	
Gln	Val	Thr	Ala	Thr	Met	Gly	Gln	Leu	Gln	Gln	Leu	Val	Ala	Ala	Gly
				325				330						335	
Gly	Gly	Pro	Ser	Gln	Leu	Ala	Ser	Met	Gly	Ser	Gln	Gln	Ala	Gln	Leu
			340					345					350		
Ile	Ser	Ser	Gln	Ala	Gln	Gln	Gly	Gly	Gln	Gln	His	Ala	Thr	Leu	Val
		355					360					365			
Ser	Asp	Lys	Lys	Glu	Asp	Glu	Glu	Gly	Val	Ala	Glu	Ala	Glu	Arg	Ala
	370					375					380				
Pro	Ile	Asp	Ala	Gly	Thr	Ala	Ala	Ser	Gln	Arg	Gly	Gln	Glu	Gly	Thr
385				390					395					400	

---

Val Leu

<210> 55

<211> 103

<212> PRT

<213> Mycobacterium tuberculosis

<220>

<223> Rv3865 - conserved hypothetical protein

<400> 55

Met	Thr	Gly	Phe	Leu	Gly	Val	Val	Pro	Ser	Phe	Leu	Lys	Val	Leu	Ala
1				5				10					15		

---

Gly Met His Asn Glu Ile Val Gly Asp Ile Lys Arg Ala Thr Asp Thr

20

25

30

Val Ala Gly Ile Ser Gly Arg Val Gln Leu Thr His Gly Ser Phe Thr  
           35                                  40                                  45

Ser Lys Phe Asn Asp Thr Leu Gln Glu Phe Glu Thr Thr Arg Ser Ser  
       50                                  55                                  60

Thr Gly Thr Gly Leu Gln Gly Val Thr Ser Gly Leu Ala Asn Asn Leu  
   65                                  70                                  75                                  80

Leu Ala Ala Ala Gly Ala Tyr Leu Lys Ala Asp Asp Gly Leu Ala Gly  
                                   85                                  90                                  95

Val Ile Asp Lys Ile Phe Gly  
                   100

<210> 56

<211> 283

<212> PRT

<213> Mycobacterium tuberculosis

<220>

<223> Rv3866 - conserved hypothetical protein

<400> 56

Met Thr Gly Pro Ser Ala Ala Gly Arg Ala Gly Thr Ala Asp Asn Val  
   1                                  5                                  10                                  15

Val Gly Val Glu Val Thr Ile Asp Gly Met Leu Val Ile Ala Asp Arg  
                   20                                  25                                  30

Leu His Leu Val Asp Phe Pro Val Thr Leu Gly Ile Arg Pro Asn Ile  
                   35                                  40                                  45

Pro Gln Glu Asp Leu Arg Asp Ile Val Trp Glu Gln Val Gln Arg Asp  
                   50                                  55                                  60

Leu Thr Ala Gln Gly Val Leu Asp Leu His Gly Glu Pro Gln Pro Thr  
   65                                  70                                  75                                  80

Val Ala Glu Met Val Glu Thr Leu Gly Arg Pro Asp Arg Thr Leu Glu  
                                   85                                  90                                  95

Gly Arg Trp Trp Arg Arg Asp Ile Gly Gly Val Met Val Arg Phe Val  
                   100                                  105                                  110

Val Cys Arg Arg Gly Asp Arg His Val Ile Ala Ala Arg Asp Gly Asp  
                   115                                  120                                  125

Met Leu Val Leu Gln Leu Val Ala Pro Gln Val Gly Leu Ala Gly Met  
                   130                                  135                                  140

Val Thr Ala Val Leu Gly Pro Ala Glu Pro Ala Asn Val Glu Pro Leu  
   145                                  150                                  155                                  160

Thr Gly Val Ala Thr Glu Leu Ala Glu Cys Thr Thr Ala Ser Gln Leu  
                                   165                                  170                                  175

Thr Gln Tyr Gly Ile Ala Pro Ala Ser Ala Arg Val Tyr Ala Glu Ile

180	185	190
Val Gly Asn Pro Thr Gly Trp	Val Glu Ile Val Ala Ser Gln Arg His	
195	200	205
Pro Gly Gly Thr Thr Thr Gln Thr Asp Ala Ala Ala Gly Val Leu Asp		
210	215	220
Ser Lys Leu Gly Arg Leu Val Ser Leu Pro Arg Arg Val Gly Gly Asp		
225	230	235
Leu Tyr Gly Ser Phe Leu Pro Gly Thr Gln Gln Asn Leu Glu Arg Ala		
245	250	255
Leu Asp Gly Leu Leu Glu Leu Leu Pro Ala Gly Ala Trp Leu Asp His		
260	265	270
Thr Ser Asp His Ala Gln Ala Ser Ser Arg Gly		
275	280	

<210> 57  
 <211> 183  
 <212> PRT  
 <213> mycobacterium tuberculosis

<220>  
 <223> Protein sequence Rv3867

<400> 57

Met Val Asp Pro Pro Gly Asn Asp Asp Asp His Gly Asp Leu Asp Ala		
1	5	10
Leu Asp Phe Ser Ala Ala His Thr Asn Glu Ala Ser Pro Leu Asp Ala		
20	25	30
Leu Asp Asp Tyr Ala Pro Val Gln Thr Asp Asp Ala Glu Gly Asp Leu		
35	40	45
Asp Ala Leu His Ala Leu Thr Glu Arg Asp Glu Glu Pro Glu Leu Glu		
50	55	60
Leu Phe Thr Val Thr Asn Pro Gln Gly Ser Val Ser Val Ser Thr Leu		
65	70	75
Met Asp Gly Arg Ile Gln His Val Glu Leu Thr Asp Lys Ala Thr Ser		
85	90	95
Met Ser Glu Ala Gln Leu Ala Asp Glu Ile Phe Val Ile Ala Asp Leu		
100	105	110
Ala Arg Gln Lys Ala Arg Ala Ser Gln Tyr Thr Phe Met Val Glu Asn		
115	120	125
Ile Gly Glu Leu Thr Asp Glu Asp Ala Glu Gly Ser Ala Leu Leu Arg		
130	135	140
Glu Phe Val Gly Met Thr Leu Asn Leu Pro Thr Pro Glu Glu Ala Ala		
145	150	155
		160

Ala Ala Glu Ala Glu Val Phe Ala Thr Arg Tyr Asp Val Asp Tyr Thr  
 165 170 175

Ser Arg Tyr Lys Ala Asp Asp  
 180

<210> 58

<211> 573

<212> PRT

<213> mycobacterium tuberculosis

<220>

<223> Protein sequence Rv3868

<400> 58

Met Thr Asp Arg Leu Ala Ser Leu Phe Glu Ser Ala Val Ser Met Leu  
 1 5 10 15

Pro Met Ser Glu Ala Arg Ser Leu Asp Leu Phe Thr Glu Ile Thr Asn  
 20 25 30

Tyr Asp Glu Ser Ala Cys Asp Ala Trp Ile Gly Arg Ile Arg Cys Gly  
 35 40 45

Asp Thr Asp Arg Val Thr Leu Phe Arg Ala Trp Tyr Ser Arg Arg Asn  
 50 55 60

Phe Gly Gln Leu Ser Gly Ser Val Gln Ile Ser Met Ser Thr Leu Asn  
 65 70 75 80

Ala Arg Ile Ala Ile Gly Gly Leu Tyr Gly Asp Ile Thr Tyr Pro Val  
 85 90 95

Thr Ser Pro Leu Ala Ile Thr Met Gly Phe Ala Ala Cys Glu Ala Ala  
 100 105 110

Gln Gly Asn Tyr Ala Asp Ala Met Glu Ala Leu Glu Ala Ala Pro Val  
 115 120 125

Ala Gly Ser Glu His Leu Val Ala Trp Met Lys Ala Val Val Tyr Gly  
 130 135 140

Ala Ala Glu Arg Trp Thr Asp Val Ile Asp Gln Val Lys Ser Ala Gly  
 145 150 155 160

Lys Trp Pro Asp Lys Phe Leu Ala Gly Ala Ala Gly Val Ala His Gly  
 165 170 175

Val Ala Ala Ala Asn Leu Ala Leu Phe Thr Glu Ala Glu Arg Arg Leu  
 180 185 190

Thr Glu Ala Asn Asp Ser Pro Ala Gly Glu Ala Cys Ala Arg Ala Ile  
 195 200 205

Ala Trp Tyr Leu Ala Met Ala Arg Arg Ser Gln Gly Asn Glu Ser Ala  
 210 215 220

Ala Val Ala Leu Leu Glu Trp Leu Gln Thr Thr His Pro Glu Pro Lys  
 225 230 235 240

Val Ala Ala Ala Leu Lys Asp Pro Ser Tyr Arg Leu Lys Thr Thr Thr



245					250					255					
Ala	Glu	Gln	Ile	Ala	Ser	Arg	Ala	Asp	Pro	Trp	Asp	Pro	Gly	Ser	Val
			260					265					270		
Val	Thr	Asp	Asn	Ser	Gly	Arg	Glu	Arg	Leu	Leu	Ala	Glu	Ala	Gln	Ala
		275					280					285			
Glu	Leu	Asp	Arg	Gln	Ile	Gly	Leu	Thr	Arg	Val	Lys	Asn	Gln	Ile	Glu
		290					295					300			
Arg	Tyr	Arg	Ala	Ala	Thr	Leu	Met	Ala	Arg	Val	Arg	Ala	Ala	Lys	Gly
							310					315			320
Met	Lys	Val	Ala	Gln	Pro	Ser	Lys	His	Met	Ile	Phe	Thr	Gly	Pro	Pro
				325					330					335	
Gly	Thr	Gly	Lys	Thr	Thr	Ile	Ala	Arg	Val	Val	Ala	Asn	Ile	Leu	Ala
			340					345					350		
Gly	Leu	Gly	Val	Ile	Ala	Glu	Pro	Lys	Leu	Val	Glu	Thr	Ser	Arg	Lys
		355					360					365			
Asp	Phe	Val	Ala	Glu	Tyr	Glu	Gly	Gln	Ser	Ala	Val	Lys	Thr	Ala	Lys
	370					375					380				
Thr	Ile	Asp	Gln	Ala	Leu	Gly	Gly	Val	Leu	Phe	Ile	Asp	Glu	Ala	Tyr
				390								395			400
Ala	Leu	Val	Gln	Glu	Arg	Asp	Gly	Arg	Thr	Asp	Pro	Phe	Gly	Gln	Glu
			405						410					415	
Ala	Leu	Asp	Thr	Leu	Leu	Ala	Arg	Met	Glu	Asn	Asp	Arg	Asp	Arg	Leu
			420					425					430		
Val	Val	Ile	Ile	Ala	Gly	Tyr	Ser	Ser	Asp	Ile	Asp	Arg	Leu	Leu	Glu
		435					440					445			
Thr	Asn	Glu	Gly	Leu	Arg	Ser	Arg	Phe	Ala	Thr	Arg	Ile	Glu	Phe	Asp
		450					455					460			
Thr	Tyr	Ser	Pro	Glu	Glu	Leu	Leu	Glu	Ile	Ala	Asn	Val	Ile	Ala	Ala
				465			470					475			480
Ala	Asp	Asp	Ser	Ala	Leu	Thr	Ala	Glu	Ala	Ala	Glu	Asn	Phe	Leu	Gln
				485				490					495		
Ala	Ala	Lys	Gln	Leu	Glu	Gln	Arg	Met	Leu	Arg	Gly	Arg	Arg	Ala	Leu
			500					505					510		
Asp	Val	Ala	Gly	Asn	Gly	Arg	Tyr	Ala	Arg	Gln	Leu	Val	Glu	Ala	Ser
		515					520					525			
Glu	Gln	Cys	Arg	Asp	Met	Arg	Leu	Ala	Gln	Val	Leu	Asp	Ile	Asp	Thr
		530					535					540			
Leu	Asp	Glu	Asp	Arg	Leu	Arg	Glu	Ile	Asn	Gly	Ser	Asp	Met	Ala	Glu
		545					550					555			560
Ala	Ile	Ala	Ala	Val	His	Ala	His	Leu	Asn	Met	Arg	Glu			
				565				570							

<210> 59  
 <211> 480  
 <212> PRT  
 <213> mycobacterium tuberculosis

<220>  
 <223> Protein sequence Rv3869

<400> 59

Met Gly Leu Arg Leu Thr Thr Lys Val Gln Val Ser Gly Trp Arg Phe  
 1 5 10 15

Leu Leu Arg Arg Leu Glu His Ala Ile Val Arg Arg Asp Thr Arg Met  
 20 25 30

Phe Asp Asp Pro Leu Gln Phe Tyr Ser Arg Ser Ile Ala Leu Gly Ile  
 35 40 45

Val Val Ala Val Leu Ile Leu Ala Gly Ala Ala Leu Leu Ala Tyr Phe  
 50 55 60

Lys Pro Gln Gly Lys Leu Gly Gly Thr Ser Leu Phe Thr Asp Arg Ala  
 65 70 75 80

Thr Asn Gln Leu Tyr Val Leu Leu Ser Gly Gln Leu His Pro Val Tyr  
 85 90 95

Asn Leu Thr Ser Ala Arg Leu Val Leu Gly Asn Pro Ala Asn Pro Ala  
 100 105 110

Thr Val Lys Ser Ser Glu Leu Ser Lys Leu Pro Met Gly Gln Thr Val  
 115 120 125

Gly Ile Pro Gly Ala Pro Tyr Ala Thr Pro Val Ser Ala Gly Ser Thr  
 130 135 140

Ser Ile Trp Thr Leu Cys Asp Thr Val Ala Arg Ala Asp Ser Thr Ser  
 145 150 155 160

Pro Val Val Gln Thr Ala Val Ile Ala Met Pro Leu Glu Ile Asp Ala  
 165 170 175

---

Ser Ile Asp Pro Leu Gln Ser His Glu Ala Val Leu Val Ser Tyr Gln  
 180 185 190

Gly Glu Thr Trp Ile Val Thr Thr Lys Gly Arg His Ala Ile Asp Leu  
 195 200 205

Thr Asp Arg Ala Leu Thr Ser Ser Met Gly Ile Pro Val Thr Ala Arg  
 210 215 220

Pro Thr Pro Ile Ser Glu Gly Met Phe Asn Ala Leu Pro Asp Met Gly  
 225 230 235 240

Pro Trp Gln Leu Pro Pro Ile Pro Ala Ala Gly Ala Pro Asn Ser Leu  
 245 250 255

---

Gly Leu Pro Asp Asp Leu Val Ile Gly Ser Val Phe Gln Ile His Thr

260					265					270					
Asp	Lys	Gly	Pro	Gln	Tyr	Tyr	Val	Val	Leu	Pro	Asp	Gly	Ile	Ala	Gln
		275					280					285			
Val	Asn	Ala	Thr	Thr	Ala	Ala	Ala	Leu	Arg	Ala	Thr	Gln	Ala	His	Gly
	290					295					300				
Leu	Val	Ala	Pro	Pro	Ala	Met	Val	Pro	Ser	Leu	Val	Val	Arg	Ile	Ala
305					310					315					320
Glu	Arg	Val	Tyr	Pro	Ser	Pro	Leu	Pro	Asp	Glu	Pro	Leu	Lys	Ile	Val
				325					330					335	
Ser	Arg	Pro	Gln	Asp	Pro	Ala	Leu	Cys	Trp	Ser	Trp	Gln	Arg	Ser	Ala
			340					345					350		
Gly	Asp	Gln	Ser	Pro	Gln	Ser	Thr	Val	Leu	Ser	Gly	Arg	His	Leu	Pro
	355						360					365			
Ile	Ser	Pro	Ser	Ala	Met	Asn	Met	Gly	Ile	Lys	Gln	Ile	His	Gly	Thr
	370					375					380				
Ala	Thr	Val	Tyr	Leu	Asp	Gly	Gly	Lys	Phe	Val	Ala	Leu	Gln	Ser	Pro
385					390					395					400
Asp	Pro	Arg	Tyr	Thr	Glu	Ser	Met	Tyr	Tyr	Ile	Asp	Pro	Gln	Gly	Val
				405					410					415	
Arg	Tyr	Gly	Val	Pro	Asn	Ala	Glu	Thr	Ala	Lys	Ser	Leu	Gly	Leu	Ser
			420					425					430		
Ser	Pro	Gln	Asn	Ala	Pro	Trp	Glu	Ile	Val	Arg	Leu	Leu	Val	Asp	Gly
		435					440					445			
Pro	Val	Leu	Ser	Lys	Asp	Ala	Ala	Leu	Leu	Glu	His	Asp	Thr	Leu	Pro
	450				455						460				
Ala	Asp	Pro	Ser	Pro	Arg	Lys	Val	Pro	Ala	Gly	Ala	Ser	Gly	Ala	Pro
465					470					475					480

&lt;210&gt; 60

&lt;211&gt; 747

&lt;212&gt; PRT

&lt;213&gt; mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; Protein sequence Rv3870

&lt;400&gt; 60

Met	Thr	Thr	Lys	Lys	Phe	Thr	Pro	Thr	Ile	Thr	Arg	Gly	Pro	Arg	Leu
1				5					10					15	
Thr	Pro	Gly	Glu	Ile	Ser	Leu	Thr	Pro	Pro	Asp	Asp	Leu	Gly	Ile	Asp
		20						25					30		
Ile	Pro	Pro	Ser	Gly	Val	Gln	Lys	Ile	Leu	Pro	Tyr	Val	Met	Gly	Gly
		35					40					45			

Ala Met Leu Gly Met Ile Ala Ile Met Val Ala Gly Gly Thr Arg Gln  
50 55 60

Leu Ser Pro Tyr Met Leu Met Met Pro Leu Met Met Ile Val Met Met  
65 70 75 80

Val Gly Gly Leu Ala Gly Ser Thr Gly Gly Gly Gly Lys Lys Val Pro  
85 90 95

Glu Ile Asn Ala Asp Arg Lys Glu Tyr Leu Arg Tyr Leu Ala Gly Leu  
100 105 110

Arg Thr Arg Val Thr Ser Ser Ala Thr Ser Gln Val Ala Phe Phe Ser  
115 120 125

Tyr His Ala Pro His Pro Glu Asp Leu Leu Ser Ile Val Gly Thr Gln  
130 135 140

Arg Gln Trp Ser Arg Pro Ala Asn Ala Asp Phe Tyr Ala Ala Thr Arg  
145 150 155 160

Ile Gly Ile Gly Asp Gln Pro Ala Val Asp Arg Leu Leu Lys Pro Ala  
165 170 175

Val Gly Gly Glu Leu Ala Ala Ala Ser Ala Ala Pro Gln Pro Phe Leu  
180 185 190

Glu Pro Val Ser His Met Trp Val Val Lys Phe Leu Arg Thr His Gly  
195 200 205

Leu Ile His Asp Cys Pro Lys Leu Leu Gln Leu Arg Thr Phe Pro Thr  
210 215 220

Ile Ala Ile Gly Gly Asp Leu Ala Gly Ala Ala Gly Leu Met Thr Ala  
225 230 235 240

Met Ile Cys His Leu Ala Val Phe His Pro Pro Asp Leu Leu Gln Ile  
245 250 255

Arg Val Leu Thr Glu Glu Pro Asp Asp Pro Asp Trp Ser Trp Leu Lys  
260 265 270

Trp Leu Pro His Val Gln His Gln Thr Glu Thr Asp Ala Ala Gly Ser  
275 280 285

Thr Arg Leu Ile Phe Thr Arg Gln Glu Gly Leu Ser Asp Leu Ala Ala  
290 295 300

Arg Gly Pro His Ala Pro Asp Ser Leu Pro Gly Gly Pro Tyr Val Val  
305 310 315 320

Val Val Asp Leu Thr Gly Gly Lys Ala Gly Phe Pro Pro Asp Gly Arg  
325 330 335

Ala Gly Val Thr Val Ile Thr Leu Gly Asn His Arg Gly Ser Ala Tyr  
340 345 350

Arg Ile Arg Val His Glu Asp Gly Thr Ala Asp Asp Arg Leu Pro Asn  
355 360 365

Gln Ser Phe Arg Gln Val Thr Ser Val Thr Asp Arg Met Ser Pro Gln

370	375	380
Gln Ala Ser Arg Ile	Ala Arg Lys Leu	Ala Gly Trp Ser Ile Thr Gly
385	390	395 400
Thr Ile Leu Asp	Lys Thr Ser Arg Val	Gln Lys Lys Val Ala Thr Asp
	405	410 415
Trp His Gln Leu Val	Gly Ala Gln Ser Val	Glu Glu Ile Thr Pro Ser
	420	425 430
Arg Trp Arg Met Tyr	Thr Asp Thr Asp Arg	Asp Arg Leu Lys Ile Pro
	435	440 445
Phe Gly His Glu Leu	Lys Thr Gly Asn Val	Met Tyr Leu Asp Ile Lys
	450	455 460
Glu Gly Ala Glu Phe	Gly Ala Gly Pro His	Gly Met Leu Ile Gly Thr
465	470	475 480
Thr Gly Ser Gly Lys	Ser Glu Phe Leu Arg	Thr Leu Ile Leu Ser Leu
	485	490 495
Val Ala Met Thr His	Pro Asp Gln Val Asn	Leu Leu Leu Thr Asp Phe
	500	505 510
Lys Gly Gly Ser Thr	Phe Leu Gly Met Glu	Lys Leu Pro His Thr Ala
	515	520 525
Ala Val Val Thr Asn	Met Ala Glu Glu Ala	Glu Leu Val Ser Arg Met
	530	535 540
Gly Glu Val Leu Thr	Gly Glu Leu Asp Arg	Arg Gln Ser Ile Leu Arg
545	550	555 560
Gln Ala Gly Met Lys	Val Gly Ala Ala Gly	Ala Leu Ser Gly Val Ala
	565	570 575
Glu Tyr Glu Lys Tyr	Arg Glu Arg Gly Ala	Asp Leu Pro Pro Leu Pro
	580	585 590
Thr Leu Phe Val Val	Val Asp Glu Phe Ala	Glu Leu Leu Gln Ser His
	595	600 605
Pro Asp Phe Ile Gly	Leu Phe Asp Arg Ile	Cys Arg Val Gly Arg Ser
610	615	620
Leu Arg Val His Leu	Leu Leu Ala Thr Gln	Ser Leu Gln Thr Gly Gly
625	630	635 640
Val Arg Ile Asp Lys	Leu Glu Pro Asn Leu	Thr Tyr Arg Ile Ala Leu
	645	650 655
Arg Thr Thr Ser Ser	His Glu Ser Lys Ala	Val Ile Gly Thr Pro Glu
	660	665 670
Ala Gln Tyr Ile Thr	Asn Lys Glu Ser Gly	Val Gly Phe Leu Arg Val
	675	680 685
Gly Met Glu Asp Pro	Val Lys Phe Ser Thr	Phe Tyr Ile Ser Gly Pro
690	695	700

Tyr Met Pro Pro Ala Ala Gly Val Glu Thr Asn Gly Glu Ala Gly Gly  
705 710 715 720

Pro Gly Gln Gln Thr Thr Arg Gln Ala Ala Arg Ile His Arg Phe Thr  
725 730 735

Ala Ala Pro Val Leu Glu Glu Ala Pro Thr Pro  
740 745

<210> 61

<211> 591

<212> PRT

<213> mycobacterium tuberculosis

<220>

<223> Protein sequence Rv3871

<400> 61

Met Thr Ala Glu Pro Glu Val Arg Thr Leu Arg Glu Val Val Leu Asp  
1 5 10 15

Gln Leu Gly Thr Ala Glu Ser Arg Ala Tyr Lys Met Trp Leu Pro Pro  
20 25 30

Leu Thr Asn Pro Val Pro Leu Asn Glu Leu Ile Ala Arg Asp Arg Arg  
35 40 45

Gln Pro Leu Arg Phe Ala Leu Gly Ile Met Asp Glu Pro Arg Arg His  
50 55 60

Leu Gln Asp Val Trp Gly Val Asp Val Ser Gly Ala Gly Gly Asn Ile  
65 70 75 80

Gly Ile Gly Gly Ala Pro Gln Thr Gly Lys Ser Thr Leu Leu Gln Thr  
85 90 95

Met Val Met Ser Ala Ala Ala Thr His Ser Pro Arg Asn Val Gln Phe  
100 105 110

Tyr Cys Ile Asp Leu Gly Gly Gly Gly Leu Ile Tyr Leu Glu Asn Leu  
115 120 125

Pro His Val Gly Gly Val Ala Asn Arg Ser Glu Pro Asp Lys Val Asn  
130 135 140

Arg Val Val Ala Glu Met Gln Ala Val Met Arg Gln Arg Glu Thr Thr  
145 150 155 160

Phe Lys Glu His Arg Val Gly Ser Ile Gly Met Tyr Arg Gln Leu Arg  
165 170 175

Asp Asp Pro Ser Gln Pro Val Ala Ser Asp Pro Tyr Gly Asp Val Phe  
180 185 190

Leu Ile Ile Asp Gly Trp Pro Gly Phe Val Gly Glu Phe Pro Asp Leu  
195 200 205

Glu Gly Gln Val Gln Asp Leu Ala Ala Gln Gly Leu Ala Phe Gly Val

210	215	220
His Val Ile Ile Ser Thr Pro Arg Trp Thr Glu Leu Lys Ser Arg Val 225 230 235 240		
Arg Asp Tyr Leu Gly Thr Lys Ile Glu Phe Arg Leu Gly Asp Val Asn 245 250 255		
Glu Thr Gln Ile Asp Arg Ile Thr Arg Glu Ile Pro Ala Asn Arg Pro 260 265 270		
Gly Arg Ala Val Ser Met Glu Lys His His Leu Met Ile Gly Val Pro 275 280 285		
Arg Phe Asp Gly Val His Ser Ala Asp Asn Leu Val Glu Ala Ile Thr 290 295 300		
Ala Gly Val Thr Gln Ile Ala Ser Gln His Thr Glu Gln Ala Pro Pro 305 310 315 320		
Val Arg Val Leu Pro Glu Arg Ile His Leu His Glu Leu Asp Pro Asn 325 330 335		
Pro Pro Gly Pro Glu Ser Asp Tyr Arg Thr Arg Trp Glu Ile Pro Ile 340 345 350		
Gly Leu Arg Glu Thr Asp Leu Thr Pro Ala His Cys His Met His Thr 355 360 365		
Asn Pro His Leu Leu Ile Phe Gly Ala Ala Lys Ser Gly Lys Thr Thr 370 375 380		
Ile Ala His Ala Ile Ala Arg Ala Ile Cys Ala Arg Asn Ser Pro Gln 385 390 395 400		
Gln Val Arg Phe Met Leu Ala Asp Tyr Arg Ser Gly Leu Leu Asp Ala 405 410 415		
Val Pro Asp Thr His Leu Leu Gly Ala Gly Ala Ile Asn Arg Asn Ser 420 425 430		
Ala Ser Leu Asp Glu Ala Val Gln Ala Leu Ala Val Asn Leu Lys Lys 435 440 445		
Arg Leu Pro Pro Thr Asp Leu Thr Thr Ala Gln Leu Arg Ser Arg Ser 450 455 460		
Trp Trp Ser Gly Phe Asp Val Val Leu Leu Val Asp Asp Trp His Met 465 470 475 480		
Ile Val Gly Ala Ala Gly Gly Met Pro Pro Met Ala Pro Leu Ala Pro 485 490 495		
Leu Leu Pro Ala Ala Ala Asp Ile Gly Leu His Ile Ile Val Thr Cys 500 505 510		
Gln Met Ser Gln Ala Tyr Lys Ala Thr Met Asp Lys Phe Val Gly Ala 515 520 525		
Ala Phe Gly Ser Gly Ala Pro Thr Met Phe Leu Ser Gly Glu Lys Gln 530 535 540		

Glu Phe Pro Ser Ser Glu Phe Lys Val Lys Arg Arg Pro Pro Gly Gln  
 545 550 555 560

Ala Phe Leu Val Ser Pro Asp Gly Lys Glu Val Ile Gln Ala Pro Tyr  
 565 570 575

Ile Glu Pro Pro Glu Glu Val Phe Ala Ala Pro Pro Ser Ala Gly  
 580 585 590

<210> 62

<211> 99

<212> PRT

<213> Mycobacterium tuberculosis

<220>

<223> Rv3872-PE35 - PE family-related protein

<400> 62

Met Glu Lys Met Ser His Asp Pro Ile Ala Ala Asp Ile Gly Thr Gln  
 1 5 10 15

Val Ser Asp Asn Ala Leu His Gly Val Thr Ala Gly Ser Thr Ala Leu  
 20 25 30

Thr Ser Val Thr Gly Leu Val Pro Ala Gly Ala Asp Glu Val Ser Ala  
 35 40 45

Gln Ala Ala Thr Ala Phe Thr Ser Glu Gly Ile Gln Leu Leu Ala Ser  
 50 55 60

Asn Ala Ser Ala Gln Asp Gln Leu His Arg Ala Gly Glu Ala Val Gln  
 65 70 75 80

Asp Val Ala Arg Thr Tyr Ser Gln Ile Asp Asp Gly Ala Ala Gly Val  
 85 90 95

Phe Ala Glu

<210> 63

<211> 368

<212> PRT

<213> Mycobacterium tuberculosis

<220>

<223> Rv3873-PPE68 - PPE family protein

<400> 63

Met Leu Trp His Ala Met Pro Pro Glu Leu Asn Thr Ala Arg Leu Met  
 1 5 10 15

Ala Gly Ala Gly Pro Ala Pro Met Leu Ala Ala Ala Ala Gly Trp Gln  
 20 25 30

Thr Leu Ser Ala Ala Leu Asp Ala Gln Ala Val Glu Leu Thr Ala Arg  
 35 40 45

Leu Asn Ser Leu Gly Glu Ala Trp Thr Gly Gly Gly Ser Asp Lys Ala



50					55					60					
Leu	Ala	Ala	Ala	Thr	Pro	Met	Val	Val	Trp	Leu	Gln	Thr	Ala	Ser	Thr
65					70					75					80
Gln	Ala	Lys	Thr	Arg	Ala	Met	Gln	Ala	Thr	Ala	Gln	Ala	Ala	Ala	Tyr
				85					90						95
Thr	Gln	Ala	Met	Ala	Thr	Thr	Pro	Ser	Leu	Pro	Glu	Ile	Ala	Ala	Asn
			100					105					110		
His	Ile	Thr	Gln	Ala	Val	Leu	Thr	Ala	Thr	Asn	Phe	Phe	Gly	Ile	Asn
			115					120					125		
Thr	Ile	Pro	Ile	Ala	Leu	Thr	Glu	Met	Asp	Tyr	Phe	Ile	Arg	Met	Trp
			130					135					140		
Asn	Gln	Ala	Ala	Leu	Ala	Met	Glu	Val	Tyr	Gln	Ala	Glu	Thr	Ala	Val
145					150					155					160
Asn	Thr	Leu	Phe	Glu	Lys	Leu	Glu	Pro	Met	Ala	Ser	Ile	Leu	Asp	Pro
				165					170					175	
Gly	Ala	Ser	Gln	Ser	Thr	Thr	Asn	Pro	Ile	Phe	Gly	Met	Pro	Ser	Pro
			180					185					190		
Gly	Ser	Ser	Thr	Pro	Val	Gly	Gln	Leu	Pro	Pro	Ala	Ala	Thr	Gln	Thr
			195					200					205		
Leu	Gly	Gln	Leu	Gly	Glu	Met	Ser	Gly	Pro	Met	Gln	Gln	Leu	Thr	Gln
			210					215					220		
Pro	Leu	Gln	Gln	Val	Thr	Ser	Leu	Phe	Ser	Gln	Val	Gly	Gly	Thr	Gly
225					230					235					240
Gly	Gly	Asn	Pro	Ala	Asp	Glu	Glu	Ala	Ala	Gln	Met	Gly	Leu	Leu	Gly
				245					250					255	
Thr	Ser	Pro	Leu	Ser	Asn	His	Pro	Leu	Ala	Gly	Gly	Ser	Gly	Pro	Ser
			260					265					270		
Ala	Gly	Ala	Gly	Leu	Leu	Arg	Ala	Glu	Ser	Leu	Pro	Gly	Ala	Gly	Gly
			275				280					285			
Ser	Leu	Thr	Arg	Thr	Pro	Leu	Met	Ser	Gln	Leu	Ile	Glu	Lys	Pro	Val
			290				295					300			
Ala	Pro	Ser	Val	Met	Pro	Ala	Ala	Ala	Ala	Gly	Ser	Ser	Ala	Thr	Gly
305					310					315					320
Gly	Ala	Ala	Pro	Val	Gly	Ala	Gly	Ala	Met	Gly	Gln	Gly	Ala	Gln	Ser
				325					330					335	
Gly	Gly	Ser	Thr	Arg	Pro	Gly	Leu	Val	Ala	Pro	Ala	Pro	Leu	Ala	Gln
			340					345					350		
Glu	Arg	Glu	Glu	Asp	Asp	Glu	Asp	Asp	Trp	Asp	Glu	Glu	Asp	Asp	Trp
			355				360					365			

&lt;210&gt; 64

&lt;211&gt; 100

&lt;212&gt; PRT

&lt;213&gt; Mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; Rv3874-esxB - 10kDa culture filtrate antigen CFP10

&lt;400&gt; 64

Met Ala Glu Met Lys Thr Asp Ala Ala Thr Leu Ala Gln Glu Ala Gly  
 1 5 10 15

Asn Phe Glu Arg Ile Ser Gly Asp Leu Lys Thr Gln Ile Asp Gln Val  
 20 25 30

Glu Ser Thr Ala Gly Ser Leu Gln Gly Gln Trp Arg Gly Ala Ala Gly  
 35 40 45

Thr Ala Ala Gln Ala Ala Val Val Arg Phe Gln Glu Ala Ala Asn Lys  
 50 55 60

Gln Lys Gln Glu Leu Asp Glu Ile Ser Thr Asn Ile Arg Gln Ala Gly  
 65 70 75 80

Val Gln Tyr Ser Arg Ala Asp Glu Glu Gln Gln Gln Ala Leu Ser Ser  
 85 90 95

Gln Met Gly Phe  
 100

&lt;210&gt; 65

&lt;211&gt; 95

&lt;212&gt; PRT

&lt;213&gt; Mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; Rv3875-Esat6 - 6 kDa early secretory antigenic target Esat6 (Esat-6)

&lt;400&gt; 65

Met Thr Glu Gln Gln Trp Asn Phe Ala Gly Ile Glu Ala Ala Ala Ser  
 1 5 10 15

Ala Ile Gln Gly Asn Val Thr Ser Ile His Ser Leu Leu Asp Glu Gly  
 20 25 30

Lys Gln Ser Leu Thr Lys Leu Ala Ala Ala Trp Gly Gly Ser Gly Ser  
 35 40 45

Glu Ala Tyr Gln Gly Val Gln Gln Lys Trp Asp Ala Thr Ala Thr Glu  
 50 55 60

Leu Asn Asn Ala Leu Gln Asn Leu Ala Arg Thr Ile Ser Glu Ala Gly  
 65 70 75 80

Gln Ala Met Ala Ser Thr Glu Gly Asn Val Thr Gly Met Phe Ala  
 85 90 95

<210> 66  
 <211> 666  
 <212> PRT  
 <213> mycobacterium tuberculosis

<220>  
 <223> Protein sequence Rv3876

<400> 66

Met	Ala	Ala	Asp	Tyr	Asp	Lys	Leu	Phe	Arg	Pro	His	Glu	Gly	Met	Glu	1	5	10	15
Ala	Pro	Asp	Asp	Met	Ala	Ala	Gln	Pro	Phe	Phe	Asp	Pro	Ser	Ala	Ser	20	25	30	
Phe	Pro	Pro	Ala	Pro	Ala	Ser	Ala	Asn	Leu	Pro	Lys	Pro	Asn	Gly	Gln	35	40	45	
Thr	Pro	Pro	Pro	Thr	Ser	Asp	Asp	Leu	Ser	Glu	Arg	Phe	Val	Ser	Ala	50	55	60	
Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Thr	Pro	Met	65	70	75	80
Pro	Ile	Ala	Ala	Gly	Glu	Pro	Pro	Ser	Pro	Glu	Pro	Ala	Ala	Ser	Lys	85	90	95	
Pro	Pro	Thr	Pro	Pro	Met	Pro	Ile	Ala	Gly	Pro	Glu	Pro	Ala	Pro	Pro	100	105	110	
Lys	Pro	Pro	Thr	Pro	Pro	Met	Pro	Ile	Ala	Gly	Pro	Glu	Pro	Ala	Pro	115	120	125	
Pro	Lys	Pro	Pro	Thr	Pro	Pro	Met	Pro	Ile	Ala	Gly	Pro	Ala	Pro	Thr	130	135	140	
Pro	Thr	Glu	Ser	Gln	Leu	Ala	Pro	Pro	Arg	Pro	Pro	Thr	Pro	Gln	Thr	145	150	155	160
Pro	Thr	Gly	Ala	Pro	Gln	Gln	Pro	Glu	Ser	Pro	Ala	Pro	His	Val	Pro	165	170	175	
Ser	His	Gly	Pro	His	Gln	Pro	Arg	Arg	Thr	Ala	Pro	Ala	Pro	Pro	Trp	180	185	190	
Ala	Lys	Met	Pro	Ile	Gly	Glu	Pro	Pro	Pro	Ala	Pro	Ser	Arg	Pro	Ser	195	200	205	
Ala	Ser	Pro	Ala	Glu	Pro	Pro	Thr	Arg	Pro	Ala	Pro	Gln	His	Ser	Arg	210	215	220	
Arg	Ala	Arg	Arg	Gly	His	Arg	Tyr	Arg	Thr	Asp	Thr	Glu	Arg	Asn	Val	225	230	235	240
Gly	Lys	Val	Ala	Thr	Gly	Pro	Ser	Ile	Gln	Ala	Arg	Leu	Arg	Ala	Glu	245	250	255	
Glu	Ala	Ser	Gly	Ala	Gln	Leu	Ala	Pro	Gly	Thr	Glu	Pro	Ser	Pro	Ala	260	265	270	

Pro Leu Gly Gln Pro Arg Ser Tyr Leu Ala Pro Pro Thr Arg Pro Ala  
275 280 285

Pro Thr Glu Pro Pro Pro Ser Pro Ser Pro Gln Arg Asn Ser Gly Arg  
290 295 300

Arg Ala Glu Arg Arg Val His Pro Asp Leu Ala Ala Gln His Ala Ala  
305 310 315 320

Ala Gln Pro Asp Ser Ile Thr Ala Ala Thr Thr Gly Gly Arg Arg Arg  
325 330 335

Lys Arg Ala Ala Pro Asp Leu Asp Ala Thr Gln Lys Ser Leu Arg Pro  
340 345 350

Ala Ala Lys Gly Pro Lys Val Lys Lys Val Lys Pro Gln Lys Pro Lys  
355 360 365

Ala Thr Lys Pro Pro Lys Val Val Ser Gln Arg Gly Trp Arg His Trp  
370 375 380

Val His Ala Leu Thr Arg Ile Asn Leu Gly Leu Ser Pro Asp Glu Lys  
385 390 395 400

Tyr Glu Leu Asp Leu His Ala Arg Val Arg Arg Asn Pro Arg Gly Ser  
405 410 415

Tyr Gln Ile Ala Val Val Gly Leu Lys Gly Gly Ala Gly Lys Thr Thr  
420 425 430

Leu Thr Ala Ala Leu Gly Ser Thr Leu Ala Gln Val Arg Ala Asp Arg  
435 440 445

Ile Leu Ala Leu Asp Ala Asp Pro Gly Ala Gly Asn Leu Ala Asp Arg  
450 455 460

Val Gly Arg Gln Ser Gly Ala Thr Ile Ala Asp Val Leu Ala Glu Lys  
465 470 475 480

Glu Leu Ser His Tyr Asn Asp Ile Arg Ala His Thr Ser Val Asn Ala  
485 490 495

Val Asn Leu Glu Val Leu Pro Ala Pro Glu Tyr Ser Ser Ala Gln Arg  
500 505 510

Ala Leu Ser Asp Ala Asp Trp His Phe Ile Ala Asp Pro Ala Ser Arg  
515 520 525

Phe Tyr Asn Leu Val Leu Ala Asp Cys Gly Ala Gly Phe Phe Asp Pro  
530 535 540

Leu Thr Arg Gly Val Leu Ser Thr Val Ser Gly Val Val Val Val Ala  
545 550 555 560

Ser Val Ser Ile Asp Gly Ala Gln Gln Ala Ser Val Ala Leu Asp Trp  
565 570 575

Leu Arg Asn Asn Gly Tyr Gln Asp Leu Ala Ser Arg Ala Cys Val Val  
580 585 590

Ile Asn His Ile Met Pro Gly Glu Pro Asn Val Ala Val Lys Asp Leu

595

600

605

Val Arg His Phe Glu Gln Gln Val Gln Pro Gly Arg Val Val Val Met  
 610 615 620

Pro Trp Asp Arg His Ile Ala Ala Gly Thr Glu Ile Ser Leu Asp Leu  
 625 630 635 640

Leu Asp Pro Ile Tyr Lys Arg Lys Val Leu Glu Leu Ala Ala Ala Leu  
 645 650 655

Ser Asp Asp Phe Glu Arg Ala Gly Arg Arg  
 660 665

<210> 67  
 <211> 511  
 <212> PRT  
 <213> mycobacterium tuberculosis

<220>  
 <223> Protein sequence Rv3877

<400> 67

Met Ser Ala Pro Ala Val Ala Ala Gly Pro Thr Ala Ala Gly Ala Thr  
 1 5 10 15

Ala Ala Arg Pro Ala Thr Thr Arg Val Thr Ile Leu Thr Gly Arg Arg  
 20 25 30

Met Thr Asp Leu Val Leu Pro Ala Ala Val Pro Met Glu Thr Tyr Ile  
 35 40 45

Asp Asp Thr Val Ala Val Leu Ser Glu Val Leu Glu Asp Thr Pro Ala  
 50 55 60

Asp Val Leu Gly Gly Phe Asp Phe Thr Ala Gln Gly Val Trp Ala Phe  
 65 70 75 80

Ala Arg Pro Gly Ser Pro Pro Leu Lys Leu Asp Gln Ser Leu Asp Asp  
 85 90 95

Ala Gly Val Val Asp Gly Ser Leu Leu Thr Leu Val Ser Val Ser Arg  
 100 105 110

---

Thr Glu Arg Tyr Arg Pro Leu Val Glu Asp Val Ile Asp Ala Ile Ala  
 115 120 125

Val Leu Asp Glu Ser Pro Glu Phe Asp Arg Thr Ala Leu Asn Arg Phe  
 130 135 140

Val Gly Ala Ala Ile Pro Leu Leu Thr Ala Pro Val Ile Gly Met Ala  
 145 150 155 160

Met Arg Ala Trp Trp Glu Thr Gly Arg Ser Leu Trp Trp Pro Leu Ala  
 165 170 175

Ile Gly Ile Leu Gly Ile Ala Val Leu Val Gly Ser Phe Val Ala Asn  
 180 185 190

---

Arg Phe Tyr Gln Ser Gly His Leu Ala Glu Cys Leu Leu Val Thr Thr  
195 200 205

Tyr Leu Leu Ile Ala Thr Ala Ala Ala Leu Ala Val Pro Leu Pro Arg  
210 215 220

Gly Val Asn Ser Leu Gly Ala Pro Gln Val Ala Gly Ala Ala Thr Ala  
225 230 235 240

Val Leu Phe Leu Thr Leu Met Thr Arg Gly Gly Pro Arg Lys Arg His  
245 250 255

Glu Leu Ala Ser Phe Ala Val Ile Thr Ala Ile Ala Val Ile Ala Ala  
260 265 270

Ala Ala Ala Phe Gly Tyr Gly Tyr Gln Asp Trp Val Pro Ala Gly Gly  
275 280 285

Ile Ala Phe Gly Leu Phe Ile Val Thr Asn Ala Ala Lys Leu Thr Val  
290 295 300

Ala Val Ala Arg Ile Ala Leu Pro Pro Ile Pro Val Pro Gly Glu Thr  
305 310 315 320

Val Asp Asn Glu Glu Leu Leu Asp Pro Val Ala Thr Pro Glu Ala Thr  
325 330 335

Ser Glu Glu Thr Pro Thr Trp Gln Ala Ile Ile Ala Ser Val Pro Ala  
340 345 350

Ser Ala Val Arg Leu Thr Glu Arg Ser Lys Leu Ala Lys Gln Leu Leu  
355 360 365

Ile Gly Tyr Val Thr Ser Gly Thr Leu Ile Leu Ala Ala Gly Ala Ile  
370 375 380

Ala Val Val Val Arg Gly His Phe Phe Val His Ser Leu Val Val Ala  
385 390 395 400

Gly Leu Ile Thr Thr Val Cys Gly Phe Arg Ser Arg Leu Tyr Ala Glu  
405 410 415

Arg Trp Cys Ala Trp Ala Leu Leu Ala Ala Thr Val Ala Ile Pro Thr  
420 425 430

Gly Leu Thr Ala Lys Leu Ile Ile Trp Tyr Pro His Tyr Ala Trp Leu  
435 440 445

Leu Leu Ser Val Tyr Leu Thr Val Ala Leu Val Ala Leu Val Val Val  
450 455 460

Gly Ser Met Ala His Val Arg Arg Val Ser Pro Val Val Lys Arg Thr  
465 470 475 480

Leu Glu Leu Ile Asp Gly Ala Met Ile Ala Ala Ile Ile Pro Met Leu  
485 490 495

Leu Trp Ile Thr Gly Val Tyr Asp Thr Val Arg Asn Ile Arg Phe  
500 505 510

<210> 68  
 <211> 280  
 <212> PRT  
 <213> Mycobacterium tuberculosis

<220>

<223> Rv3878 - conserved hypothetical alanine rich protein

<400> 68

Met Ala Glu Pro Leu Ala Val Asp Pro Thr Gly Leu Ser Ala Ala Ala  
 1 5 10 15

Ala Lys Leu Ala Gly Leu Val Phe Pro Gln Pro Pro Ala Pro Ile Ala  
 20 25 30

Val Ser Gly Thr Asp Ser Val Val Ala Ala Ile Asn Glu Thr Met Pro  
 35 40 45

Ser Ile Glu Ser Leu Val Ser Asp Gly Leu Pro Gly Val Lys Ala Ala  
 50 55 60

Leu Thr Arg Thr Ala Ser Asn Met Asn Ala Ala Ala Asp Val Tyr Ala  
 65 70 75 80

Lys Thr Asp Gln Ser Leu Gly Thr Ser Leu Ser Gln Tyr Ala Phe Gly  
 85 90 95

Ser Ser Gly Glu Gly Leu Ala Gly Val Ala Ser Val Gly Gly Gln Pro  
 100 105 110

Ser Gln Ala Thr Gln Leu Leu Ser Thr Pro Val Ser Gln Val Thr Thr  
 115 120 125

Gln Leu Gly Glu Thr Ala Ala Glu Leu Ala Pro Arg Val Val Ala Thr  
 130 135 140

Val Pro Gln Leu Val Gln Leu Ala Pro His Ala Val Gln Met Ser Gln  
 145 150 155 160

Asn Ala Ser Pro Ile Ala Gln Thr Ile Ser Gln Thr Ala Gln Gln Ala  
 165 170 175

Ala Gln Ser Ala Gln Gly Gly Ser Gly Pro Met Pro Ala Gln Leu Ala  
 180 185 190

Ser Ala Glu Lys Pro Ala Thr Glu Gln Ala Glu Pro Val His Glu Val  
 195 200 205

Thr Asn Asp Asp Gln Gly Asp Gln Gly Asp Val Gln Pro Ala Glu Val  
 210 215 220

Val Ala Ala Ala Arg Asp Glu Gly Ala Gly Ala Ser Pro Gly Gln Gln  
 225 230 235 240

Pro Gly Gly Gly Val Pro Ala Gln Ala Met Asp Thr Gly Ala Gly Ala  
 245 250 255

Arg Pro Ala Ala Ser Pro Leu Ala Ala Pro Val Asp Pro Ser Thr Pro  
 260 265 270

Ala Pro Ser Thr Thr Thr Thr Leu  
275 280

<210> 69

<211> 729

<212> PRT

<213> Mycobacterium tuberculosis

<220>

<223> Rv3879c - hypothetical alanine and proline rich  
protein

<400> 69

Met Ser Ile Thr Arg Pro Thr Gly Ser Tyr Ala Arg Gln Met Leu Asp  
1 5 10 15

Pro Gly Gly Trp Val Glu Ala Asp Glu Asp Thr Phe Tyr Asp Arg Ala  
20 25 30

Gln Glu Tyr Ser Gln Val Leu Gln Arg Val Thr Asp Val Leu Asp Thr  
35 40 45

Cys Arg Gln Gln Lys Gly His Val Phe Glu Gly Gly Leu Trp Ser Gly  
50 55 60

Gly Ala Ala Asn Ala Ala Asn Gly Ala Leu Gly Ala Asn Ile Asn Gln  
65 70 75 80

Leu Met Thr Leu Gln Asp Tyr Leu Ala Thr Val Ile Thr Trp His Arg  
85 90 95

His Ile Ala Gly Leu Ile Glu Gln Ala Lys Ser Asp Ile Gly Asn Asn  
100 105 110

Val Asp Gly Ala Gln Arg Glu Ile Asp Ile Leu Glu Asn Asp Pro Ser  
115 120 125

Leu Asp Ala Asp Glu Arg His Thr Ala Ile Asn Ser Leu Val Thr Ala  
130 135 140

Thr His Gly Ala Asn Val Ser Leu Val Ala Glu Thr Ala Glu Arg Val  
145 150 155 160

Leu Glu Ser Lys Asn Trp Lys Pro Pro Lys Asn Ala Leu Glu Asp Leu  
165 170 175

Leu Gln Gln Lys Ser Pro Pro Pro Pro Asp Val Pro Thr Leu Val Val  
180 185 190

Pro Ser Pro Gly Thr Pro Gly Thr Pro Gly Thr Pro Ile Thr Pro Gly  
195 200 205

Thr Pro Ile Thr Pro Gly Thr Pro Ile Thr Pro Ile Pro Gly Ala Pro  
210 215 220

Val Thr Pro Ile Thr Pro Thr Pro Gly Thr Pro Val Thr Pro Val Thr  
225 230 235 240

Pro Gly Lys Pro Val Thr Pro Val Thr Pro Val Lys Pro Gly Thr Pro  
245 250 255



Gly Glu Pro Thr Pro Ile Thr Pro Val Thr Pro Pro Val Ala Pro Ala  
260 265 270

Thr Pro Ala Thr Pro Ala Thr Pro Val Thr Pro Ala Pro Ala Pro His  
275 280 285

Pro Gln Pro Ala Pro Ala Pro Ala Pro Ser Pro Gly Pro Gln Pro Val  
290 295 300

Thr Pro Ala Thr Pro Gly Pro Ser Gly Pro Ala Thr Pro Gly Thr Pro  
305 310 315 320

Gly Gly Glu Pro Ala Pro His Val Lys Pro Ala Ala Leu Ala Glu Gln  
325 330 335

Pro Gly Val Pro Gly Gln His Ala Gly Gly Gly Thr Gln Ser Gly Pro  
340 345 350

Ala His Ala Asp Glu Ser Ala Ala Ser Val Thr Pro Ala Ala Ala Ser  
355 360 365

Gly Val Pro Gly Ala Arg Ala Ala Ala Ala Pro Ser Gly Thr Ala  
370 375 380

Val Gly Ala Gly Ala Arg Ser Ser Val Gly Thr Ala Ala Ala Ser Gly  
385 390 395 400

Ala Gly Ser His Ala Ala Thr Gly Arg Ala Pro Val Ala Thr Ser Asp  
405 410 415

Lys Ala Ala Ala Pro Ser Thr Arg Ala Ala Ser Ala Arg Thr Ala Pro  
420 425 430

Pro Ala Arg Pro Pro Ser Thr Asp His Ile Asp Lys Pro Asp Arg Ser  
435 440 445

Glu Ser Ala Asp Asp Gly Thr Pro Val Ser Met Ile Pro Val Ser Ala  
450 455 460

Ala Arg Ala Ala Arg Asp Ala Ala Thr Ala Ala Ala Ser Ala Arg Gln  
465 470 475 480

Arg Gly Arg Gly Asp Ala Leu Arg Leu Ala Arg Arg Ile Ala Ala Ala  
485 490 495

Leu Asn Ala Ser Asp Asn Asn Ala Gly Asp Tyr Gly Phe Phe Trp Ile  
500 505 510

Thr Ala Val Thr Thr Asp Gly Ser Ile Val Val Ala Asn Ser Tyr Gly  
515 520 525

Leu Ala Tyr Ile Pro Asp Gly Met Glu Leu Pro Asn Lys Val Tyr Leu  
530 535 540

Ala Ser Ala Asp His Ala Ile Pro Val Asp Glu Ile Ala Arg Cys Ala  
545 550 555 560

Thr Tyr Pro Val Leu Ala Val Gln Ala Trp Ala Ala Phe His Asp Met  
565 570 575

Thr Leu Arg Ala Val Ile Gly Thr Ala Glu Gln Leu Ala Ser Ser Asp  
 580 585 590  
 Pro Gly Val Ala Lys Ile Val Leu Glu Pro Asp Asp Ile Pro Glu Ser  
 595 600 605  
 Gly Lys Met Thr Gly Arg Ser Arg Leu Glu Val Val Asp Pro Ser Ala  
 610 615 620  
 Ala Ala Gln Leu Ala Asp Thr Thr Asp Gln Arg Leu Leu Asp Leu Leu  
 625 630 635 640  
 Pro Pro Ala Pro Val Asp Val Asn Pro Pro Gly Asp Glu Arg His Met  
 645 650 655  
 Leu Trp Phe Glu Leu Met Lys Pro Met Thr Ser Thr Ala Thr Gly Arg  
 660 665 670  
 Glu Ala Ala His Leu Arg Ala Phe Arg Ala Tyr Ala Ala His Ser Gln  
 675 680 685  
 Glu Ile Ala Leu His Gln Ala His Thr Ala Thr Asp Ala Ala Val Gln  
 690 695 700  
 Arg Val Ala Val Ala Asp Trp Leu Tyr Trp Gln Tyr Val Thr Gly Leu  
 705 710 715 720  
 Leu Asp Arg Ala Leu Ala Ala Ala Cys  
 725

<210> 70

<211> 115

<212> PRT

<213> Mycobacterium tuberculosis

<220>

<223> Rv3880c - conserved hypothetical protein

<400> 70

Val Ser Met Asp Glu Leu Asp Pro His Val Ala Arg Ala Leu Thr Leu  
 1 5 10 15

Ala Ala Arg Phe Gln Ser Ala Leu Asp Gly Thr Leu Asn Gln Met Asn  
 20 25 30

Asn Gly Ser Phe Arg Ala Thr Asp Glu Ala Glu Thr Val Glu Val Thr  
 35 40 45

Ile Asn Gly His Gln Trp Leu Thr Gly Leu Arg Ile Glu Asp Gly Leu  
 50 55 60

Leu Lys Lys Leu Gly Ala Glu Ala Val Ala Gln Arg Val Asn Glu Ala  
 65 70 75 80

Leu His Asn Ala Gln Ala Ala Ala Ser Ala Tyr Asn Asp Ala Ala Gly  
 85 90 95

Glu Gln Leu Thr Ala Ala Leu Ser Ala Met Ser Arg Ala Met Asn Glu  
 100 105 110

Gly Met Ala  
115

<210> 71

<211> 460

<212> PRT

<213> Mycobacterium tuberculosis

<220>

<223> Rv3881c - conserved hypothetical alanine and  
glycine rich protein

<400> 71

Met Thr Gln Ser Gln Thr Val Thr Val Asp Gln Gln Glu Ile Leu Asn  
1 5 10 15

Arg Ala Asn Glu Val Glu Ala Pro Met Ala Asp Pro Pro Thr Asp Val  
20 25 30

Pro Ile Thr Pro Cys Glu Leu Thr Ala Ala Lys Asn Ala Ala Gln Gln  
35 40 45

Leu Val Leu Ser Ala Asp Asn Met Arg Glu Tyr Leu Ala Ala Gly Ala  
50 55 60

Lys Glu Arg Gln Arg Leu Ala Thr Ser Leu Arg Asn Ala Ala Lys Ala  
65 70 75 80

Tyr Gly Glu Val Asp Glu Glu Ala Ala Thr Ala Leu Asp Asn Asp Gly  
85 90 95

Glu Gly Thr Val Gln Ala Glu Ser Ala Gly Ala Val Gly Gly Asp Ser  
100 105 110

Ser Ala Glu Leu Thr Asp Thr Pro Arg Val Ala Thr Ala Gly Glu Pro  
115 120 125

Asn Phe Met Asp Leu Lys Glu Ala Ala Arg Lys Leu Glu Thr Gly Asp  
130 135 140

Gln Gly Ala Ser Leu Ala His Phe Ala Asp Gly Trp Asn Thr Phe Asn  
145 150 155 160

Leu Thr Leu Gln Gly Asp Val Lys Arg Phe Arg Gly Phe Asp Asn Trp  
165 170 175

Glu Gly Asp Ala Ala Thr Ala Cys Glu Ala Ser Leu Asp Gln Gln Arg  
180 185 190

Gln Trp Ile Leu His Met Ala Lys Leu Ser Ala Ala Met Ala Lys Gln  
195 200 205

Ala Gln Tyr Val Ala Gln Leu His Val Trp Ala Arg Arg Glu His Pro  
210 215 220

Thr Tyr Glu Asp Ile Val Gly Leu Glu Arg Leu Tyr Ala Glu Asn Pro  
225 230 235 240

Ser Ala Arg Asp Gln Ile Leu Pro Val Tyr Ala Glu Tyr Gln Gln Arg  
245 250 255

Ser Glu Lys Val Leu Thr Glu Tyr Asn Asn Lys Ala Ala Leu Glu Pro  
260 265 270

Val Asn Pro Pro Lys Pro Pro Pro Ala Ile Lys Ile Asp Pro Pro Pro  
275 280 285

Pro Pro Gln Glu Gln Gly Leu Ile Pro Gly Phe Leu Met Pro Pro Ser  
290 295 300

Asp Gly Ser Gly Val Thr Pro Gly Thr Gly Met Pro Ala Ala Pro Met  
305 310 315 320

Val Pro Pro Thr Gly Ser Pro Gly Gly Gly Leu Pro Ala Asp Thr Ala  
325 330 335

Ala Gln Leu Thr Ser Ala Gly Arg Glu Ala Ala Ala Leu Ser Gly Asp  
340 345 350

Val Ala Val Lys Ala Ala Ser Leu Gly Gly Gly Gly Gly Gly Gly Val  
355 360 365

Pro Ser Ala Pro Leu Gly Ser Ala Ile Gly Gly Ala Glu Ser Val Arg  
370 375 380

Pro Ala Gly Ala Gly Asp Ile Ala Gly Leu Gly Gln Gly Arg Ala Gly  
385 390 395 400

Gly Gly Ala Ala Leu Gly Gly Gly Gly Gly Met Gly Met Pro Met Gly Ala  
405 410 415

Ala His Gln Gly Gln Gly Gly Ala Lys Ser Lys Gly Ser Gln Gln Glu  
420 425 430

Asp Glu Ala Leu Tyr Thr Glu Asp Arg Ala Trp Thr Glu Ala Val Ile  
435 440 445

Gly Asn Arg Arg Arg Gln Asp Ser Lys Glu Ser Lys  
450 455 460

<210> 72

<211> 462

<212> PRT

<213> *Mycobacterium tuberculosis*

<220>

<223> Rv3882c - possible conserved membrane protein

<400> 72

Met Arg Asn Pro Leu Gly Leu Arg Phe Ser Thr Gly His Ala Leu Leu  
1 5 10 15

Ala Ser Ala Leu Ala Pro Pro Cys Ile Ile Ala Phe Leu Glu Thr Arg  
20 25 30

Tyr Trp Trp Ala Gly Ile Ala Leu Ala Ser Leu Gly Val Ile Val Ala  
35 40 45

Thr Val Thr Phe Tyr Gly Arg Arg Ile Thr Gly Trp Val Ala Ala Val  
50 55 60

Tyr	Ala	Trp	Leu	Arg	Arg	Arg	Arg	Arg	Pro	Pro	Asp	Ser	Ser	Ser	Glu	65	70	75	80
Pro	Val	Val	Gly	Ala	Thr	Val	Lys	Pro	Gly	Asp	His	Val	Ala	Val	Arg	85	90	95	
Trp	Gln	Gly	Glu	Phe	Leu	Val	Ala	Val	Ile	Glu	Leu	Ile	Pro	Arg	Pro	100	105	110	
Phe	Thr	Pro	Thr	Val	Ile	Val	Asp	Gly	Gln	Ala	His	Thr	Asp	Asp	Met	115	120	125	
Leu	Asp	Thr	Gly	Leu	Val	Glu	Glu	Leu	Leu	Ser	Val	His	Cys	Pro	Asp	130	135	140	
Leu	Glu	Ala	Asp	Ile	Val	Ser	Ala	Gly	Tyr	Arg	Val	Gly	Asn	Thr	Ala	145	150	155	160
Ala	Pro	Asp	Val	Val	Ser	Leu	Tyr	Gln	Gln	Val	Ile	Gly	Thr	Asp	Pro	165	170	175	
Ala	Pro	Ala	Asn	Arg	Arg	Thr	Trp	Ile	Val	Leu	Arg	Ala	Asp	Pro	Glu	180	185	190	
Arg	Thr	Arg	Lys	Ser	Ala	Gln	Arg	Arg	Asp	Glu	Gly	Val	Ala	Gly	Leu	195	200	205	
Ala	Arg	Tyr	Leu	Val	Ala	Ser	Ala	Thr	Arg	Ile	Ala	Asp	Arg	Leu	Ala	210	215	220	
Ser	His	Gly	Val	Asp	Ala	Val	Cys	Gly	Arg	Ser	Phe	Asp	Asp	Tyr	Asp	225	230	235	240
His	Ala	Thr	Asp	Ile	Gly	Phe	Val	Arg	Glu	Lys	Trp	Ser	Met	Ile	Lys	245	250	255	
Gly	Arg	Asp	Ala	Tyr	Thr	Ala	Ala	Tyr	Ala	Ala	Pro	Gly	Gly	Pro	Asp	260	265	270	
Val	Trp	Trp	Ser	Ala	Arg	Ala	Asp	His	Thr	Ile	Thr	Arg	Val	Arg	Val	275	280	285	
Ala	Pro	Gly	Met	Ala	Pro	Gln	Ser	Thr	Val	Leu	Leu	Thr	Thr	Ala	Asp	290	295	300	
Lys	Pro	Lys	Thr	Pro	Arg	Gly	Phe	Ala	Arg	Leu	Phe	Gly	Gly	Gln	Arg	305	310	315	320
Pro	Ala	Leu	Gln	Gly	Gln	His	Leu	Val	Ala	Asn	Arg	His	Cys	Gln	Leu	325	330	335	
Pro	Ile	Gly	Ser	Ala	Gly	Val	Leu	Val	Gly	Glu	Thr	Val	Asn	Arg	Cys	340	345	350	
Pro	Val	Tyr	Met	Pro	Phe	Asp	Asp	Val	Asp	Ile	Ala	Leu	Asn	Leu	Gly	355	360	365	
Asp	Ala	Gln	Thr	Phe	Thr	Gln	Phe	Val	Val	Arg	Ala	Ala	Ala	Ala	Gly	370	375	380	

Ala Met Val Thr Val Gly Pro Gln Phe Glu Glu Phe Ala Arg Leu Ile  
385 390 395 400

Gly Ala His Ile Gly Gln Glu Val Lys Val Ala Trp Pro Asn Ala Thr  
405 410 415

Thr Tyr Leu Gly Pro His Pro Gly Ile Asp Arg Val Ile Leu Arg His  
420 425 430

Asn Val Ile Gly Thr Pro Arg His Arg Gln Leu Pro Ile Arg Arg Val  
435 440 445

Ser Pro Pro Glu Glu Ser Arg Tyr Gln Met Ala Leu Pro Lys  
450 455 460

<210> 73

<211> 446

<212> PRT

<213> Mycobacterium tuberculosis

<220>

<223> Rv3883c - possible secreted protease

<400> 73

Val His Arg Ile Phe Leu Ile Thr Val Ala Leu Ala Leu Leu Thr Ala  
1 5 10 15

Ser Pro Ala Ser Ala Ile Thr Pro Pro Pro Ile Asp Pro Gly Ala Leu  
20 25 30

Pro Pro Asp Val Thr Gly Pro Asp Gln Pro Thr Glu Gln Arg Val Leu  
35 40 45

Cys Ala Ser Pro Thr Thr Leu Pro Gly Ser Gly Phe His Asp Pro Pro  
50 55 60

Trp Ser Asn Thr Tyr Leu Gly Val Ala Asp Ala His Lys Phe Ala Thr  
65 70 75 80

Gly Ala Gly Val Thr Val Ala Val Ile Asp Thr Gly Val Asp Ala Ser  
85 90 95

Pro Arg Val Pro Ala Glu Pro Gly Gly Asp Phe Val Asp Gln Ala Gly  
100 105 110

Asn Gly Leu Ser Asp Cys Asp Ala His Gly Thr Leu Thr Ala Ser Ile  
115 120 125

Ile Ala Gly Arg Pro Ala Pro Thr Asp Gly Phe Val Gly Val Ala Pro  
130 135 140

Asp Ala Arg Leu Leu Ser Leu Arg Gln Thr Ser Glu Ala Phe Glu Pro  
145 150 155 160

Val Gly Ser Gln Ala Asn Pro Asn Asp Pro Asn Ala Thr Pro Ala Ala  
165 170 175

Gly Ser Ile Arg Ser Leu Ala Arg Ala Val Val His Ala Ala Asn Leu  
180 185 190

Gly	Val	Gly	Val	Ile	Asn	Ile	Ser	Glu	Ala	Ala	Cys	Tyr	Lys	Val	Ser
	195						200					205			
Arg	Pro	Ile	Asp	Glu	Thr	Ser	Leu	Gly	Ala	Ser	Ile	Asp	Tyr	Ala	Val
	210					215					220				
Asn	Val	Lys	Gly	Val	Val	Val	Val	Val	Ala	Ala	Gly	Asn	Thr	Gly	Gly
225				230					235						240
Asp	Cys	Val	Gln	Asn	Pro	Ala	Pro	Asp	Pro	Ser	Thr	Pro	Gly	Asp	Pro
			245					250						255	
Arg	Gly	Trp	Asn	Asn	Val	Gln	Thr	Val	Val	Thr	Pro	Ala	Trp	Tyr	Ala
			260					265					270		
Pro	Leu	Val	Leu	Ser	Val	Gly	Gly	Ile	Gly	Gln	Thr	Gly	Met	Pro	Ser
	275						280					285			
Ser	Phe	Ser	Met	His	Gly	Pro	Trp	Val	Asp	Val	Ala	Ala	Pro	Ala	Glu
	290					295					300				
Asn	Ile	Val	Ala	Leu	Gly	Asp	Thr	Gly	Glu	Pro	Val	Asn	Ala	Leu	Gln
305					310					315					320
Gly	Arg	Glu	Gly	Pro	Val	Pro	Ile	Ala	Gly	Thr	Ser	Phe	Ala	Ala	Ala
				325					330					335	
Tyr	Val	Ser	Gly	Leu	Ala	Ala	Leu	Leu	Arg	Gln	Arg	Phe	Pro	Asp	Leu
			340					345					350		
Thr	Pro	Ala	Gln	Ile	Ile	His	Arg	Ile	Thr	Ala	Thr	Ala	Arg	His	Pro
		355					360					365			
Gly	Gly	Gly	Val	Asp	Asp	Leu	Val	Gly	Ala	Gly	Val	Ile	Asp	Ala	Val
	370					375					380				
Ala	Ala	Leu	Thr	Trp	Asp	Ile	Pro	Pro	Gly	Pro	Ala	Ser	Ala	Pro	Tyr
385					390					395					400
Asn	Val	Arg	Arg	Leu	Pro	Pro	Pro	Val	Val	Glu	Pro	Gly	Pro	Asp	Arg
				405					410					415	
Arg	Pro	Ile	Thr	Ala	Val	Ala	Leu	Val	Ala	Val	Gly	Leu	Thr	Leu	Ala
			420					425					430		

---

Leu	Gly	Leu	Gly	Ala	Leu	Ala	Arg	Arg	Ala	Leu	Ser	Arg	Arg
	435						440					445	

&lt;210&gt; 74

&lt;211&gt; 619

&lt;212&gt; PRT

&lt;213&gt; Mycobacterium tuberculosis

&lt;220&gt;

&lt;223&gt; Rv3884c - probable CBXX/CFQX family protein

&lt;400&gt; 74

Met	Ser	Arg	Met	Val	Asp	Thr	Met	Gly	Asp	Leu	Leu	Thr	Ala	Arg	Arg
1				5				10						15	

His	Phe	Asp	Arg	Ala	Met	Thr	Ile	Lys	Asn	Gly	Gln	Gly	Cys	Val	Ala		
			20					25					30				
Ala	Leu	Pro	Glu	Phe	Val	Ala	Ala	Thr	Glu	Ala	Asp	Pro	Ser	Met	Ala		
		35					40					45					
Asp	Ala	Trp	Leu	Gly	Arg	Ile	Ala	Cys	Gly	Asp	Arg	Asp	Leu	Ala	Ser		
	50					55					60						
Leu	Lys	Gln	Leu	Asn	Ala	His	Ser	Glu	Trp	Leu	His	Arg	Glu	Thr	Thr		
65				70						75					80		
Arg	Ile	Gly	Arg	Thr	Leu	Ala	Ala	Glu	Val	Gln	Leu	Gly	Pro	Ser	Ile		
			85						90					95			
Gly	Ile	Thr	Val	Thr	Asp	Ala	Ser	Gln	Val	Gly	Leu	Ala	Leu	Ser	Ser		
			100					105					110				
Ala	Leu	Thr	Ile	Ala	Gly	Glu	Tyr	Ala	Lys	Ala	Asp	Ala	Leu	Leu	Ala		
		115					120					125					
Asn	Arg	Glu	Leu	Leu	Asp	Ser	Trp	Arg	Asn	Tyr	Gln	Trp	His	Gln	Leu		
	130					135					140						
Ala	Arg	Ala	Phe	Leu	Met	Tyr	Val	Thr	Gln	Arg	Trp	Pro	Asp	Val	Leu		
145				150						155					160		
Ser	Thr	Ala	Ala	Glu	Asp	Leu	Pro	Pro	Gln	Ala	Ile	Val	Met	Pro	Ala		
			165						170					175			
Val	Thr	Ala	Ser	Ile	Cys	Ala	Leu	Ala	Ala	His	Ala	Ala	Ala	His	Leu		
			180					185					190				
Gly	Gln	Gly	Arg	Val	Ala	Leu	Asp	Trp	Leu	Asp	Arg	Val	Asp	Val	Ile		
	195						200					205					
Gly	His	Ser	Arg	Ser	Ser	Glu	Arg	Phe	Gly	Ala	Asp	Val	Leu	Thr	Ala		
	210					215					220						
Ala	Ile	Gly	Pro	Ala	Asp	Ile	Pro	Leu	Leu	Val	Ala	Asp	Leu	Ala	Tyr		
225					230					235					240		
Val	Arg	Gly	Met	Val	Tyr	Arg	Gln	Leu	His	Glu	Glu	Asp	Lys	Ala	Gln		
			245						250					255			
<hr/>																	
Ile	Trp	Leu	Ser	Lys	Ala	Thr	Ile	Asn	Gly	Val	Leu	Thr	Asp	Ala	Ala		
			260					265					270				
Lys	Glu	Ala	Leu	Ala	Asp	Pro	Asn	Leu	Arg	Leu	Ile	Val	Thr	Asp	Glu		
	275						280					285					
Arg	Thr	Ile	Ala	Ser	Arg	Ser	Asp	Arg	Trp	Asp	Ala	Ser	Thr	Ala	Lys		
	290					295					300						
Ser	Arg	Asp	Gln	Leu	Asp	Asp	Asp	Asn	Ala	Ala	Gln	Arg	Arg	Gly	Glu		
305					310					315					320		
Leu	Leu	Ala	Glu	Gly	Arg	Glu	Leu	Leu	Ala	Lys	Gln	Val	Gly	Leu	Ala		
			325						330					335			
<hr/>																	
Ala	Val	Lys	Gln	Ala	Val	Ser	Ala	Leu	Glu	Asp	Gln	Leu	Glu	Val	Arg		



<211> 537

<213> Myc

 $\langle 220 \rangle$ 

<223> Rv3885c - possible conserved membrane protein

&lt;400&gt; 75

Leu Thr Ser Lys Leu Thr Gly Phe Ser Pro Arg Ser Ala Arg Arg Val  
 1 5 10 15

Ala Gly Val Trp Thr Val Phe Val Leu Ala Ser Ala Gly Trp Ala Leu  
 20 25 30

Gly Gly Gln Leu Gly Ala Val Met Ala Val Val Val Gly Val Ala Leu  
 35 40 45

Val Phe Val Gln Trp Trp Gly Gln Pro Ala Trp Ser Trp Ala Val Leu  
 50 55 60

Gly Leu Arg Gly Arg Arg Pro Val Lys Trp Asn Asp Pro Ile Thr Leu  
 65 70 75 80

Ala Asn Asn Arg Ser Gly Gly Gly Val Arg Val Gln Asp Gly Val Ala  
 85 90 95

Val Val Ala Val Gln Leu Leu Gly Arg Ala His Arg Ala Thr Thr Val  
 100 105 110

Thr Gly Ser Val Thr Val Glu Ser Asp Asn Val Ile Asp Val Val Glu  
 115 120 125

Leu Ala Pro Leu Leu Arg His Pro Leu Asp Leu Glu Leu Asp Ser Ile  
 130 135 140

Ser Val Val Thr Phe Gly Ser Arg Thr Gly Thr Val Gly Asp Tyr Pro  
 145 150 155 160

Arg Val Tyr Asp Ala Glu Ile Gly Thr Pro Pro Tyr Ala Gly Arg Arg  
 165 170 175

Glu Thr Trp Leu Ile Met Arg Leu Pro Val Ile Gly Asn Thr Gln Ala  
 180 185 190

Leu Arg Trp Arg Thr Ser Val Gly Ala Ala Ala Ile Ser Val Ala Gln  
 195 200 205

Arg Val Ala Ser Ser Leu Arg Cys Gln Gly Leu Arg Ala Lys Leu Ala  
 210 215 220

Thr Ala Thr Asp Leu Ala Glu Leu Asp Arg Arg Leu Gly Ser Asp Ala  
 225 230 235 240

Val Ala Gly Ser Ala Gln Arg Trp Lys Ala Ile Arg Gly Glu Ala Gly  
 245 250 255

Trp Met Thr Thr Tyr Ala Tyr Pro Ala Glu Ala Ile Ser Ser Arg Val  
 260 265 270

Leu Ser Gln Ala Trp Thr Leu Arg Ala Asp Glu Val Ile Gln Asn Val  
 275 280 285

Thr Val Tyr Pro Asp Ala Thr Cys Thr Ala Thr Ile Thr Val Arg Thr  
 290 295 300

Pro Thr Pro Ala Pro Thr Pro Pro Ser Val Ile Leu Arg Arg Leu Asn  
 305 310 315 320

Gly Glu Gln Ala Ala Ala Ala Ala Ala Asn Met Cys Gly Pro Arg Pro  
 325 330 335  
 His Leu Arg Gly Gln Arg Arg Cys Pro Leu Pro Ala Gln Leu Val Thr  
 340 345 350  
 Glu Ile Gly Pro Ser Gly Val Leu Ile Gly Lys Leu Ser Asn Gly Asp  
 355 360 365  
 Arg Leu Met Ile Pro Val Thr Asp Ala Gly Glu Leu Ser Arg Val Phe  
 370 375 380  
 Val Ala Ala Asp Asp Thr Ile Ala Lys Arg Ile Val Ile Arg Val Val  
 385 390 395 400  
 Gly Ala Gly Glu Arg Val Cys Val His Thr Arg Asp Gln Glu Arg Trp  
 405 410 415  
 Ala Ser Val Arg Met Pro Gln Leu Ser Ile Val Gly Thr Pro Arg Pro  
 420 425 430  
 Ala Pro Arg Thr Thr Val Gly Val Val Glu Tyr Val Arg Arg Arg Lys  
 435 440 445  
 Asn Gly Asp Asp Gly Lys Ser Glu Gly Ser Gly Val Asp Val Ala Ile  
 450 455 460  
 Ser Pro Thr Pro Arg Pro Ala Ser Val Ile Thr Ile Ala Arg Pro Gly  
 465 470 475 480  
 Thr Ser Leu Ser Glu Ser Asp Arg His Gly Phe Glu Val Thr Ile Glu  
 485 490 495  
 Gln Ile Asp Arg Ala Thr Val Lys Val Gly Ala Ala Gly Gln Asn Trp  
 500 505 510  
 Leu Val Glu Met Glu Met Phe Arg Ala Glu Asn Arg Tyr Val Ser Leu  
 515 520 525  
 Glu Pro Val Thr Met Ser Ile Gly Arg  
 530 535

---